November 13, 1961

Memorandum to: Mr. J. E. Bartelt

Subject: IBM's Relationship with X3.2

Several informal discussions with members of x3.2 were quite disquieting to me. It appears that the members are upset and disturbed about IBM's attitude toward the work of x3.2. Many members are dubious about our corporate position and feel that IBM is not acting in good faith. In particular they are extremely skeptical about IBM's assurances that we will implement or support the code. They feel that the "official" position has varied from time to time, particularly between Irv Liggett's tenure and the present day. I have tried to reassure them but have only partially succeeded.

I would like to emphasize that, in my opinion, drastic action is required. Frankly, the group feels that the standard can be reached without IBM's active support, and because of the unanimity of this feeling, I tend to agree. This looks like a potential case of IBM lagging the field (as one member stated, "as usual"). I feel you should press for implementation.

One final remark is necessary in my own defense. The committee seems to feel that my personal contributions have been important, but unnecessarily restricted by a vascialating official IBM position.

H. J. Smith

HJS/lmg
CC: Mr. W. E. Andrus, Jr.

NMAH 310/2

November 13, 1961

Memorandum to: Mr. R. W. Berner

Subject: X3.2 Code, Your Memorandum of
November 7 to Mr. J. E. Bartelt

At the November 8, 9, and 10 meeting of the x3.2 the assignment of graphics shown in the attached chart was quickly reached. The Teletype Corporation did not press for placement of the / and - in the stick containing the numeric digits, a condition which would have simplified their keyboard design. IBM presented 4 possible arrangements for characters in the 6 low order positions of the first two graphic columns and the set shown represented an acceptable arrangement for all concerned.

Later in the meeting, however, a suggested revision in placement of the control characters almost forced repositioning of the data delimiters into the first control stick. I was able to demonstrate that several alternate solutions existed, and that the suggested replacement of the data delimiters would introduce problems in deriving a 6 bit processing set with delimiters. The suggested [illegible]. It is interesting to note that if the special characters had been located in the position to which you so strongly objected, this discussion of the placement of delimiters would not have arisen, and a better overall compromise could have been achieved.

Thank you for your suggestions. They are always most welcome.

H. J. Smith

HJS/lmg
attachment

CC: Mr. W.E. Andrus, Jr.
Mr. J. E. Bartelt

NMAH 310/2

November 15, 1961

RECOMMENDED CHARACTER SET
FOR 4-ROW TELETYPEWRITERS

The attached Table 1 contains the recommended character set for Bell System 4-row teletypewriters. This is the code which has been submitted to ASA Sectional Committee X3 by the X3-2 Subcommittee as a "Proposed American Standard Code for Information Interchange." This code differs from the preliminary recommendation of October 4, 1961 in the following areas:

1. The area reserved for data delimiters has been shifted from the first eight positions of the 001 column to the last eight positions of this column. This change will permit the data processors to collapse the 7-bit set into a 6-bit set containing data delimiters, replacing only the least significant symbols (@, \ , [,] , #).

2. As a result of the above change, all of the other controls except WRU and BELL were reassigned (NULL remains in the 0000 position of the 000 column by definition). The changes are as follows:

    a. The auxiliary machine controls (R1ON, R2ON, XON, R1OF, R2OF, XOFF*) are now placed in the area formerly occupied by the data delimiters, and are arranged so that a simpler test will establish whether a given code is an auxiliary machine control and whether it is a machine activate or deactivate code. These auxiliary machine controls were previously located on five of the six alpha keys (K, L, M, N and O) which already have symbols on the upper portion of the keytop.

    b. The switching controls (SOM, EOA, EOM and EOT*) have been added and assigned so as to be on the keytops A, S, W and E which are clustered at the left of the keyboard.

    c. Horizontal tabulation, vertical tabulation and form feed have been reassigned to facilitate the collapse of the 7-bit code into a 6-bit code for communication printers.

    d. The remaining controls RU and MOTOR OFF were assigned to two of the four remaining control codes, as indicated in Table 1.

* See Legend (Section 5) of the attached "Proposed American Standard Code for Information Interchange."

NMAH 310/2

TABLE I

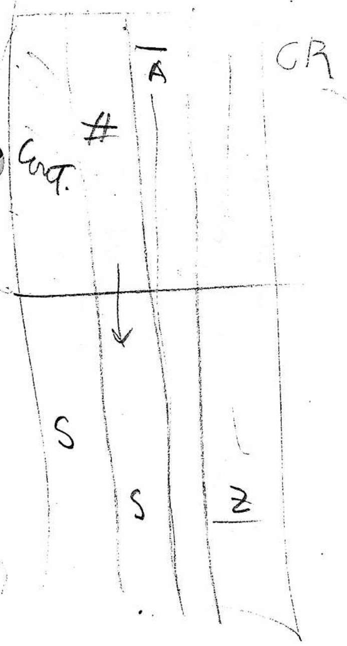
RECOMMENDED CHARACTER SET FOR 4-ROW TELETYPEWRITERS

"CONTROL" KEY

SHIFT SHIFT

b7				0	0	0	0	1	1	1	1
b6				0	0	1	1	0	0	1	1
b5				0	1	0	1	0	1	0	1
b4	b3	b2	b1								
0	0	0	0	NULL	RION	b̸	0	'	P		
		0	1	SOM	R2ON	!	1	A	Q	U	U
		1	0	MOFF	XON	"	2	B	R	N	N
		1	1	RU	EOA	<	3	C	S	A	A
0	1	0	0	WRU	RIOF	>	4	D	T	S	S
		0	1	EOT	R2OF	$	5	E	U	S	S
		1	0	HTAB	XOFF	%	6	F	V	I	I
		1	1	FF	EOM	&	7	G	W	G	G
1	0	0	0	VTAB	DATA DELIMITERS	(	8	H	X	N	N
		0	1			)	9	I	Y	E	E
		1	0	BELL		.	↑	J	Z	D	D
		1	1	x		,	←	K	@		
1	1	0	0	CR		*	+	L	\		ACK
		0	1	LF		:	;	M	[		
		1	0	UC		-	=	N	]		ESC
1	1	1	1	LC		/	?	O	#		DEL
✓	✓	✓	✓			GRAPHICS					

STUNTS | SYMBOLS | ALPHA



NOTE: SEE LEGEND (SECTION 5) OF THE ATTACHED "PROPOSED AMERICAN STANDARD CODE FOR INFORMATION INTERCHANGE".

4-Row Monocase and Up-Low Teletypewriters

November 21, 1961

MEMORANDUM:

The author reviewed the 4-row monocase and up-low teletypewriters, in terms of the recommended character set for 4-row teletypewriters, at the Teletype Corporation on November 13 and 14, 1961. This memorandum is to record the discussions.

4-Row Monocase Teletypewriter

1. The keyboard of the 33 and 35 type teletypewriters are illustrated in Figures 1 and 2 respectively. The two keyboards are identical except for the following differences necessitated by the equipment limitations and cost considerations:

	Location With Respect to Alpha Field	
Key	33 Type	35 Type
HERE IS	Right	Right side panel
CONTROL	Right	Left
BREAK	Right	Left Side panel
REPEAT	Right	Right Front
ESCAPE	Note 1	Note 2
Local Line Feed	Note 3	Left Front
Local Reverse Line Feed	Note 3	Left Front
Local Carriage Return	Not Provided	Right Front
Local Backspace	Not Provided	Right Front(ASR only)

Note 1 - The 33 type does not provide a hole in the keyboard guild plate for the ESCAPE key.

Note 2 - The 35 type can be equipped with the ESCAPE key by removing cap and inserting the ESCAPE keylever.

Note 3- The end of the platen of the 33 type protrudes through the cover to permit the attendant to rotate the platen.

2. An interlock will be provided on both the 33 and 35 types to prevent the operation of keys which do not have upper symbols, during the depression of the SHIFT key. Table 1 illustrates the conditions which would result if this interlock were not provided.

3. Inadvertent operation of the CONTROL key, simultaneously with a key not associated with a control, will result in the conditions illustrated in Table 2. The execution of the control function in the 33 and 35 types does not lend itself to a simple blocking of the non-control keys.

NMAH 310/2

4-Row Up/Low Teletypewriters

1. The 33 and 35 types will require major redesign to operate as up/low teletypewriters. Before proceeding with any development, the following are some of the items which should be resolved:

   a) Will the Bell System offer an up/low teletypewriter? If so, when?

   b) Should the development effort to increase the speed of the heavy duty teletypewriter be expanded in scope to include up/low operation?

   c) Will there be both heavy duty and light duty up/low teletypewriters?

   d) Should the lower case of the alphabet have discrete code assignments or should an up/low teletypewriter operate with a mode shift?

2. The basic code structure for an up/low teletypewriter is as follows:

000	001	010	011	100	101	110	111
	SYMBOLS	A	Z		DIGITS	a	z

This arrangement does not conform to the ASA X3-2 code structure. Therefore, to utilize the ASA code and to permit 4-row monocase and up/low teletypewriters to communicate, the up/low typing unit will have to be more complex (logically). Figure 3 illustrates possible 4-row monocase and up/low type box configurations.

3. The symbols associated with K, L, M, N, O and P must be placed on separate keys of the up/low keyboard if the SHIFT key is to be used to change b_6 from 1 to 0 to shift from the lower case of the alphabet to the upper case and to change b_5 from 1 to 0 (or 0 to 1) to shift from the lower symbols to the upper symbols (see Figure 4).

An alternate method would be to add a key, designated UPPER SYMBOL, to be depressed with keys containing upper symbol thereby retaining the monocase keyboard arrangement. This has the obvious disadvantage of differing from the "standard" typewriter mode of operation.

JBBooth

J. B. Booth

TABLE 2

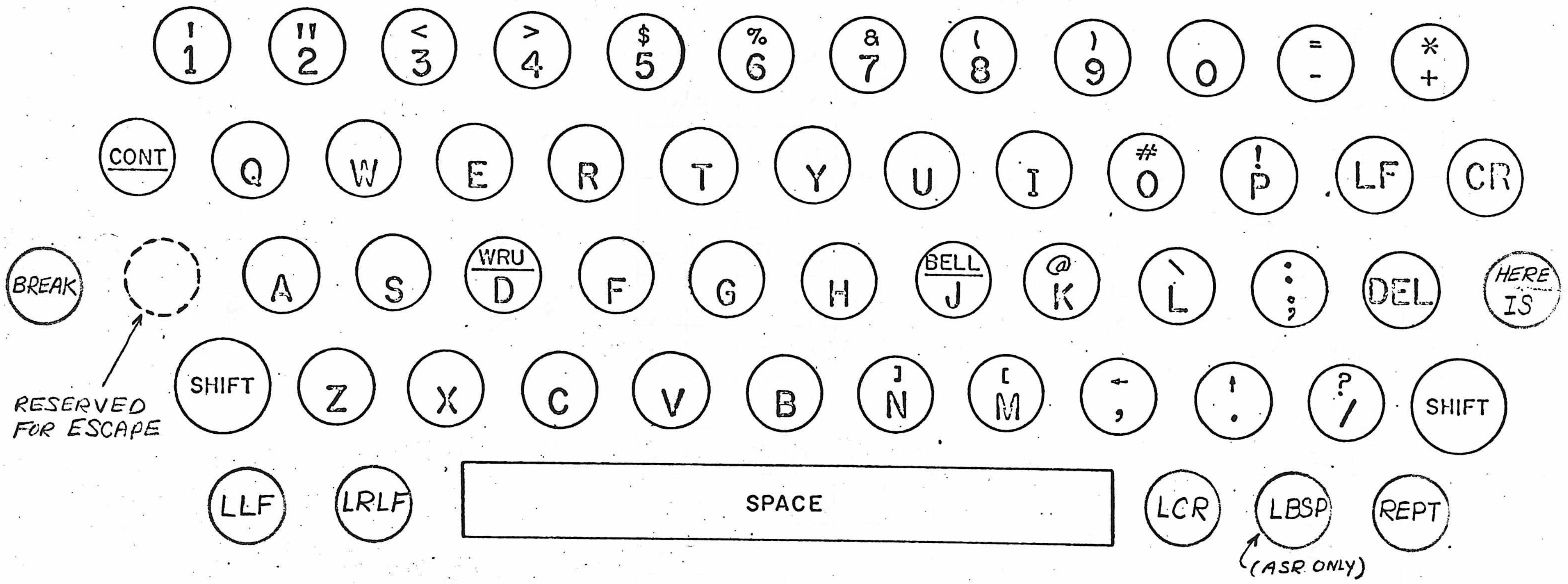
Characters Generated if Non-Control Key is Depressed Simultaneously With the CONTROL Key.

Key Depressed	Character Generated
ESCAPE	=
DELETE	?
X	Data Delimiter
Y	Data Delimiter
Z	Data Delimiter
I	(None)
K	(None)
L	Carriage Return
M	Line Feed
N	Upper Case
O	Lower Case

TABLE 3

Characters Generated if the SHIFT and CONTROL Keys are Depressed Simultaneously With the K, L, M, N, O or P Keys.

Key Depressed	Character Generated
K	Data Delimiter
L	Data Delimiter
M	Data Delimiter
N	Data Delimiter
O	Data Delimiter
P	Null



RESERVED FOR ESCAPE

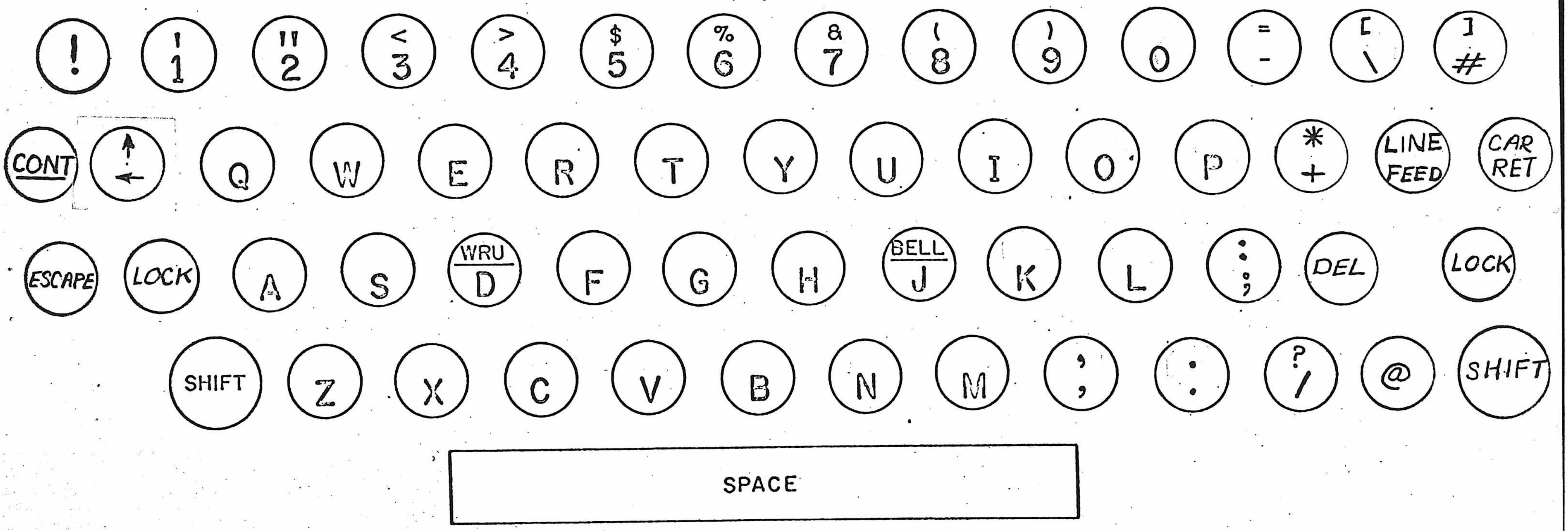
(ASR ONLY)

BASIC 4-ROW (BIT-INVERSION) KEYBOARD
(35 TYPE)

FIGURE 2

NOTE: BREAK AND
HERE IS KEYS
ARE LOCATED
ON THE KEYBOARD
SIDE PANELS.

IBR 11/15/61



Possible 4-Row Up/Low Keyboard

NOTE: AUXILIARY KEYS
ARE NOT SHOWN.

FIGURE 4

Data Processing Group

bema

Business Equipment Manufacturers Association
235 East 42nd Street, New York 17, New York — MUrray Hill 7-5969

November 28, 1961

MEMORANDUM TO: All Members X3.2

SUBJECT: Proposed American Standard Code
for Information Interchange

We are enclosing a copy of the recommended standard which your Subcommittee agreed upon at its last meeting. Copies of this recommendation are being mailed this week to all members of ASA Sectional Committee X3 and the DPG Engineering Committee.

G. M. Wilson

G. M. Wilson for
J. W. Barker, Acting Chairman X3

GMW/mh
enc.

NMAH
310/2

Formerly Office Equipment Manufacturers Institute

PROPOSED AMERICAN STANDARD CODE
FOR INFORMATION INTERCHANGE

FOREWORD

(This Foreword is not a part of the American Standard Code for Information Interchange).

Presented herein is the standard coded character set to be used for information interchange among information processing systems, communication systems, and associated equipment.

Subsequent standards will prescribe the means of implementing this standard in the principal media, such as perforated tape, punched cards and magnetic tape. Other standards will deal with collating and error control considerations. These standards will facilitate the interchange of digital information.

The 7-bit coded character set was developed by a group of highly qualified and experienced specialists in information processing and communication. Historical work in the field was reviewed, and a comprehensive program of original research and code design completed. Careful consideration has been given to the several conflicting code set requirements, and their resolution achieved in the standard code.

This standard was approved as American Standard by the American Standards Association on

Suggestions for improvement gained in the use of this standard will be welcome. They should be sent to the American Standards Association, Incorporated, 10 East 40th Street, New York 16, New York.

TABLE OF CONTENTS

Section	Description
1.	Scope
2.	Table No. 1
3.	Notations
4.	Remarks
5.	Legend
Appendixes	

PROPOSED AMERICAN STANDARD CODE
FOR INFORMATION INTERCHANGE

1. SCOPE

This coded character set is to be used for the interchange of information among information processing systems, communication systems and associated equipment.

The standard is presented with appendices on code design considerations and criteria, related subsets, extensions and deviations.

2. TABLE #1

			b_7 →	0	0	0	0	1	1	1	1
			b_6 →	0	0	1	1	0	0	1	1
			b_5 →	0	1	0	1	0	1	0	1
b_4	b_3	b_2	b_1								
0	0	0	0	NULL	R1 ON	b̸	0	!	P	↑	↑
0	0	0	1	SOM	R2 ON	'	1	A	Q		U
0	0	1	0	M OFF	X ON	"	2	B	R		N
0	0	1	1	RU	EOA	<	3	C	S	U	A
0	1	0	0	WRU	R1 OF	>	4	D	T	N	S
0	1	0	1	EOT	R2 OF	$	5	E	U	A	S
0	1	1	0	H TAB	X OFF	%	6	F	V	S	I
0	1	1	1	FF	EOM	&	7	G	W	S	G
1	0	0	0	V TAB	DD	(	8	H	X	I	N
1	0	0	1	①	DD	)	9	I	Y	G	E
1	0	1	0	BELL	DD	.	↑	J	Z	N	D
1	0	1	1	①	DD	,	←	K	@	E	↓
1	1	0	0	CR	DD	*	+	L	\	D	ACK
1	1	0	1	LF	DD	:	;	M	[		①
1	1	1	0	UC	DD	—	=	N	]		ESC
1	1	1	1	LC	DD	/	?	O	#	↓	DEL

GRAPHICS

① – Unassigned Controls

DD – Reserved for Data Delimiter

3. NOTATION

Standard 7-Bit Set code positional order and notation is shown below with b_7 the high-order and b_1 the low order bit position.

	b_7	b_6	b_5	b_4	b_3	b_2	b_1
Example - The code for "S" is	1	0	1	0	0	1	1

4. REMARKS

This Standard does not define the means by which the coded set is to be recorded in any physical medium. The standard code does not include any redundancy nor define techniques for error control. Further, it does not specify a standard collating sequence.

Deviations from the standard may create serious difficulties in information interchange and should be used only with full cognizance by the parties involved.

Unassigned codes are reserved for future standardization. Their use in information interchange prior to such standardization is a deviation from the Standard.

---

5. *LEGEND*

NULL - Null/Idle

SOM - Start of Message

M OFF - Motor off

RU - "Are you...?"

WRU - "Who are you?"

EOT - End of transmission

H TAB - Horizontal tabulation

FF - Form feed

V TAB - Vertical tabulation

BELL - Audible signal

CR - Carriage return

LF - Line feed

UC - Upper case

LC - Lower case

R1 ON - Receiver #1 on

R2 ON - Receiver #2 on

X ON - Transmitter on

EOA - End of Address

R1 OF - Receiver #1 off

R2 OF - Receiver #2 off

X OFF - Transmitter off

EOM - End of message

DD - Data delimiter

b̸ - word separator (space, normally non-printing)

< - Less than

> - Greater than

↑ - Up arrow, exponentiation

← - Left arrow, (implies, replaced by)

\ - Reverse slant

ACK - Acknowledge

ESC - Escape

DEL - Delete/Idle

NB. Expanded definitions of some of the above terms may be found in appendices.

POUGHKEEPSIE
South Road Laboratory
Dept. 272 Ext. 5541M

December 26, 1961

Memorandum To: Members of the Character Sets Committee

Subject: Meeting of the Character Sets Committee, December 19 and 20

Section 1

The meeting was held in New York City at the New Time-Life Building. The members who attended, as well as some visitors, are indicated on the attached list. For the information of members who wish to circulate information there are two methods available. Firstly, the information can be sent to Mr. C. E. Mackenzie, who will arrange to have it duplicated and forwarded; alternatively, members can send memos to other members, and in this event, should send two copies to each name on the attached list. It was agreed that the next meeting of the Committee would take place in New York City at the New Time-Life Building on Tuesday and Wednesday, January 23 and 24.

Section II

Mr. C. E. Mackenzie gave a progress report on the IBM Historic Code. He indicated that agreement to implement a code on the console typewriter of the 1410 had been obtained in the Data Systems Division. The actual details and schedules for the implementation have not been yet worked out, but this will take place rapidly.

Section III

The meeting was then, and for the remainder of the time devoted to a discussion of the X3.2 Information Interchange Standards. Mr. L. Robinson, of Data Processing, remarked that at a meeting of the Corporate Standards Technical Committee, Data Processing had agreed to this code only on the condition that it should not be implemented directly into hardware. For example, the suggestion that this code could be implemented directly in an 8-bit machine, with the realization that the IBM coding sequence had been violated, would be completely unacceptable to Data Processing. It was pointed out at this time and many other times during the meeting that the implementation of this Standard did not necessarily involve a direct use of the bit code but could certainly involve rearrangement of the bit code to incorporate the IBM collating sequence. Mr. H. J. Smith, Jr. pointed out that the standard has been very carefully worded so that it specifically excluded at this time any description of means of implementation of the code in physical media such as punched cards, magnetic tape, and perforated tape, and in CPU's. He also pointed out that task groups had been set up by the X3.2 Subcommittee to begin work on these subjects.

NMAH 310/2

Section IV

The Committee then began to discuss the possible implementation of the code in punched cards. There are three alternatives.

A. Punch the code directly into the cards. This approach is completely unacceptable to IBM. It would require the development and marketing of a new line of equipment such as keypunches, sorters, collators, and interpreters, and there is certainly no reason at this time for doing such a thing. Also, there is considerable question as to the strength of cards which may have as many as 8 punches in a column and 80 columns across the card. Product Test experience indicates that cards laced by such codes tend to buckle and jam.

B. A second approach is an expanded Hollerith Code. Mr. B. M. Updike pointed out that by considering all possible conbinations of three punches per column, two punches per column, one punch per column, and no punches per column, a total number of 299 combinations are available. If this approach were taken, it is clear that 128 combinations as in a 7-bit code, or 256 combinations as in an 8-bit code could clearly be encompassed, but not directly. If such cards could be read into an 8-bit machine, for example, translation would have to take place, either software or hardware. An advantage of this approach is that the present IBM Historic Code could be embodied directly and unchanged.

C. Some kind of compromise code can be developed. Suggestions were not forthcoming.

Section V

Four decisions were made by the committee.

A. It was agreed that before IBM could determine its official position on implementation or nonimplementation of the X3.2 Code, many factors would have to be considered and analyzed, both economic and technological. In order to obtain this information, a process very similar to that of a Product Forecasting exercise in being set in action. Under the co-ordination of Mr. John Bartelt, the Product Line Divisions will be asked to provide assumptions along the usual lines; that is to say, technology, cost, schedules, competitive situations, etc. These assumptions will be sent on to Product Forecasting, who must provide a man to work on the forecast. Marketing information will be provided by Data Processing, possibly resulting from a survey. Recognizing that this is a major effort, and that objections will probably arise as regards providing personnel, it was, nevertheless, felt that this task could be completed within three months. Three months was established as a target date. It is the responsibility of the divisional representatives on the Character Sets Committee to alert people in their own division that this effort is commencing, and that they are available to assist.

B. It is very important and in fact necessary for the X3.2 Subcommittee task groups on implementation to recognize that users are as strongly affected by the possible implementation of this code as the manufacturers. Mr. H. J. Smith, Jr. will make the point at the next meeting of the X3.2 Subcommittee that users' groups should be represented on the task groups which are going to deal with implementation in the area of punched cards, magnetic tape, perforated tape, etc.

C. In order to obtain active participation of users' groups on the X3.2 Subcommittee, Mr. John Bartelt will speak to Mr. Stan Closman about presentations on the X3.2 Information Interchange Code to both Share and Guide at their next meetings. The presentations will point out some of the impact of implementation of the code and Share and Guide will be asked to provide representation for the X3.2 Subcommittee.

D. Mr. H. J. Smith, Jr. will carry to the next meeting of the X3.2 Subcommittee the statement that the IBM Character Sets Committee does not object to the code as an information interchange code, but that no official IBM statement on implementation will be possible for a period of time. Clearly, it will be necessary for the task groups to come forth with the specific suggestions before IBM can consider the merits of such suggestions.

Section VI

Five meetings will take place in the near future relative to the work of this Committee.

A. The IBM Corporate Standards Committee meets on January 9.

B. The Engineering Committee of BEMA meets on January 10.

C. The X3 Committee meets on January 11. (In connection with this meeting, Mr. Bob Bemer should be instructed to indicate that no decision should be taken by the X3 Committee until IBM has had a chance to establish its official position in respect to the X3.2 Information Interchange Code.)

D. The X3.2 Subcommittee will meet on January 17 and 18.

E. The IBM Character Sets Committee will meet on January 23 and 24.

Section VII

Mr. W. Eadie from IBM United Kingdom reported on a code developed by the British Standards Institute which is under active consideration by ECMA. Mr. John Bartelt will distribute copies of this code to members of the Character Sets Committee.

C. E. Mackenzie
Systems Standards

CEM:jf

cc: Attached

Name	Business Address	Home Address
Mr. F. V. Adams	Dept. 411 Communications Systems Center IBM Corporation 10301 Westlake Drive Rockville, Maryland Ext. 3129	6912 Bradley Boulevard Bethesda 14, Maryland
Mr. J. E. Bartelt	GRP Staff - Standards (1-738) IBM Corporation 150 Grand Street White Plains, New York WH 9-1900 Ext. 5260	48 Richbell Road White Plains, New York 914-RO1-4644
Mr. F. P. Corr	IBM Corporation France LaGaude A. M., France	90 Ave. Ziem, Cagnes Sur Mer, A. M., France
Mr. C. J. Davis	ET Division IBM Corporation New Circle Road Lexington, Kentucky 2-2250 Ext. 3024	348 Greenbriar Road Lexington, Kentucky 7-5129
Mr. W. Eadie	IBM United Kingdom, Ltd. 101 Wigmore Street London, W. I., England	"Milverton", West Malling Kent, England
Mr. C. E. Mackenzie	Dept. 272 Bldg. 965 IBM Corporation Poughkeepsie, New York GL 4-1000 Ext. 5541M	52 Roosevelt Road Hyde Park, New York CA 9-8435
Mr. L. Robinson	IBM Corporation 112 East Post Road White Plains, New York Ext. 5068	4 Lorraine Place Scarsdale, New York
Mr. W. M. Shenko	Applied Programming Time - Life Bldg. New York, New York	144 West 16 Street New York 11, New York
Mr. H. J. Smith, Jr. ✓	Dept. 1-738 Floor 3 IBM Corporation 222 Mamaroneck Ave. White Plains, New York	50 Chadwick Place Glen Rock, New Jersey
Mr. B. M. Updike	GP Product Development Lab IBM Corporation Dept. 299 Ext. 4576 Endicott, New York	415 Crescent Place Endwell, New York PI 8-4008

PROPOSED AMERICAN STANDARD CODE FOR INFORMATION INTERCHANGE

UNOFFICIAL

January 11, 1962

NMAH 310/2

Foreword

(This Foreword is not a part of Proposed American Standard Code for Information Interchange.)

? the

This American Standard presents a standard coded character set to be used for information interchange among information processing systems, communications systems, and associated equipment.

Code design considerations and criteria, related subsets, extensions, and deviations are given in the appendixes. ? ?

Subsequent standards will prescribe the means of implementing this standard in the principal media, such as perforated tape, punched cards, and magnetic tape. Other standards will deal will collating and error control considerations. These standards will facilitate the interchange of digital information.

with

The 7-bit coded character set was developed by a group of specialists highly qualified and experienced in information processing and communication. Historical work in the field was reviewed, and a comprehensive program of original research and code design completed. Careful consideration has been given to the several conflicting code-set requirements with the purpose, achieved by this standard, of resolving these conflicts.

No

This standard was approved as American Standard by the ASA on

Suggestions for improvement gained in the use of this standard will be welcome. They should be sent to the American Standards Association, Incorporated, 10 East 40th Street, New York 16, N.Y.

Contents

Section

1. Scope
2. Standard Code
3. Positional Order and Notation
4. Legend

Appendixes

Proposed American Standard Code
for Information Interchange

1. Scope

1.1 This standard establishes a 7-bit coded character set for the interchange of data in an information processing system. No

1.2 This standard does not define the means by which the coded set is to be recorded in any physical medium. The standard code does not include any redundancy or define techniques for error control. Further, it does not specify a standard collating sequence. No

2. Standard Code

See table and notes below

b_7 →				0	0	0	0	1	1	1	1
b_6 →				0	0	1	1	0	0	1	1
b_5 →				0	1	0	1	0	1	0	1
b_4	b_3	b_2	b_1								
0	0	0	0	NULL	R1 ON	␢	0	!	P	UNASSIGNED	UNASSIGNED
0	0	0	1	SOM	R2 ON	'	1	A	Q		
0	0	1	0	M OFF	X ON	"	2	B	R		
0	0	1	1	RU	EOA	<	3	C	S		
0	1	0	0	WRU	R1 OF	>	4	D	T		
0	1	0	1	EOT	R2 OF	$	5	E	U		
0	1	1	0	H TAB	X OFF	%	6	F	V		
0	1	1	1	FF	EOM	&	7	G	W		
1	0	0	0	V TAB	DD	(	8	H	X		
1	0	0	1	Ⓘ	DD	)	9	I	Y		
1	0	1	0	BELL	DD	.	↑	J	Z		
1	0	1	1	Ⓘ	DD	,	←	K	@		
1	1	0	0	CR	DD	*	+	L	\		ACK
1	1	0	1	LF	DD	:	;	M	[		Ⓘ
1	1	1	0	UC	DD	—	=	N	]		ESC
1	1	1	1	LC	DD	/	?	O	#		DEL
						GRAPHICS					

Ⓘ – Unassigned Controls
DD – Reserved for Data Delimiter

Z @ O \

Note 1: Deviations from the standard may create serious difficulties in information interchange and should be used only with full cognizance of the parties involved.

Note 2: Unassigned codes are reserved for future standardization. Their use in information interchange prior to such standardization is a deviation from the standard.

3. Positional Order and Notation

Standard 7-bit set code positional order and notation are shown below with b_7, the high-order, and b_1, the low-order, bit position.

$$b_7\ b_6\ b_5\ b_4\ b_3\ b_2\ b_1$$

Example: The code "S" is 1 0 1 0 0 1 1

4. Legend

NULL	Null/Idle
SOM	Start of message
M off	Motor off
RU	"Are you...?"
WRU	"Who are you?"
EOT	End of transmission
H tab	Horizontal tabulation
FF	Form feed
V tab	Vertical tabulation
①	Unassigned controls
BELL	Audible signal
CR	Carriage return
LF	Line feed
UC	Upper case
LC	Lower case
R1 on	Receiver No. 1 on
R2 on	Receiver No. 2 on
X on	Transmitter on
EOA	End of address
R1 of	Receiver No. 1 off
R2 of	Receiver No. 2 off
X off	Transmitter off
EOM	End of message
DD	Data delimiter

ƀ	Word separator (space, normally non-printing)
<	Less than
>	Greater than
↑	Exponentiation
←	Replaced by
\	
ACK	Acknowledge
ESC	Escape
DEL	Delete/Idle

NOTE: Expanded definitions of some of the above terms may be found in the appendixe

December 1, 1961

TO MEMBERS OF X3.2

An emergency meeting of X3.2 is called for December 11, 12 and 13 in Washington D. C. This action is necessary so that the Appendices A, B and C can be completed and approved in time for submission to X3 sufficiently for the next X3 meeting. It is and has been our aim that these Appendices should be of high quality reflecting careful and professional work. Comments received to date indicate that there are sufficient revisions and corrections to warrant a working meeting to correct and revise said Appendices.

The meeting place will be:

TELETYPE CORPORATION OFFICE
Room 723 - Pennsylvania Bldg.
425 13th Street N. W.
Washington 4, D. C.

TEL: MEtropolitan 8-1016

In order to avoid possible conflicts and make most evident use of everyone's time, the following dates, times and strategic agenda has been set.

Dec. 11 -	7:30 P.M.	Review of Appendices and corrections and revisions which have been submitted. Establishment of working groups.
Dec. 12 -	8 to 5 P.M.	Working group meetings on Appendices.
	7:30 P.M.	Formal meeting to review progress of the working groups.
Dec. 13 -	To be scheduled at meeting of Dec. 11th.	

It is believed that this work and these meetings are extremely important for the time of completion of the present standards activities. I request that you each make every reasonable effort to attend.

Best regards,

C. E. Macon

C. E. Macon
Chairman X3.2

CEM:aks

NMAH 310/2

NMAH 310/2

461/HMcGR/CH

6th December, 1961

Mr. E.L. Tholstrup,
Friden, Inc.,
97 Humboldt Street,
ROCHESTER 2, N.Y.
U. S. A.

Dear Mr. Tholstrup:

- ASA X3.2 [illegible] -

I have now had the opportunity of discussing, with some others particularly interested in coding, the latest X3.2 proposals which you made available on an informal basis. We were very appreciative that you took the opportunity of your visit to London for the [illegible] Conference to keep us in the picture in this matter.

In order to follow up your suggestion to discuss these proposals with others active in this area, they were brought before the EEA Coding Working Party, at the suggestion of its chairman. The ASA code chart was marked "Strictly Confidential". A meeting had been arranged for December 4, and there were present:

Mr. J.H.F. Smith, Chairman - British Telecommunications Research Ltd.
Mr. H.G. Cluff - International Computers & Tabulators Ltd.
Mr. L.R. Crawley - Standard Telephones & Cables Ltd.
Mr. V. Eadie - IBM United Kingdom Ltd.
Mr. A.M. Petty - Elliott Brothers (London) Ltd.
Mr. A.J. Bagnal - E.M.I. Electronics Ltd.
Mr. J.F. Richardson - Technical Secretary, Electronic Engineering Association
Mr. H. McG. Ross - Ferranti Ltd.

All the above (except Mr. Richardson) are also members of the BSI DPE/1 Coding Committee.

Several points arose which we felt we would like to be put to the ASA X3 Committee, and I was given the responsibility of writing to you - thus I attempt to express the views of the members of the EEA Coding Working Party, without this being an official EEA statement.

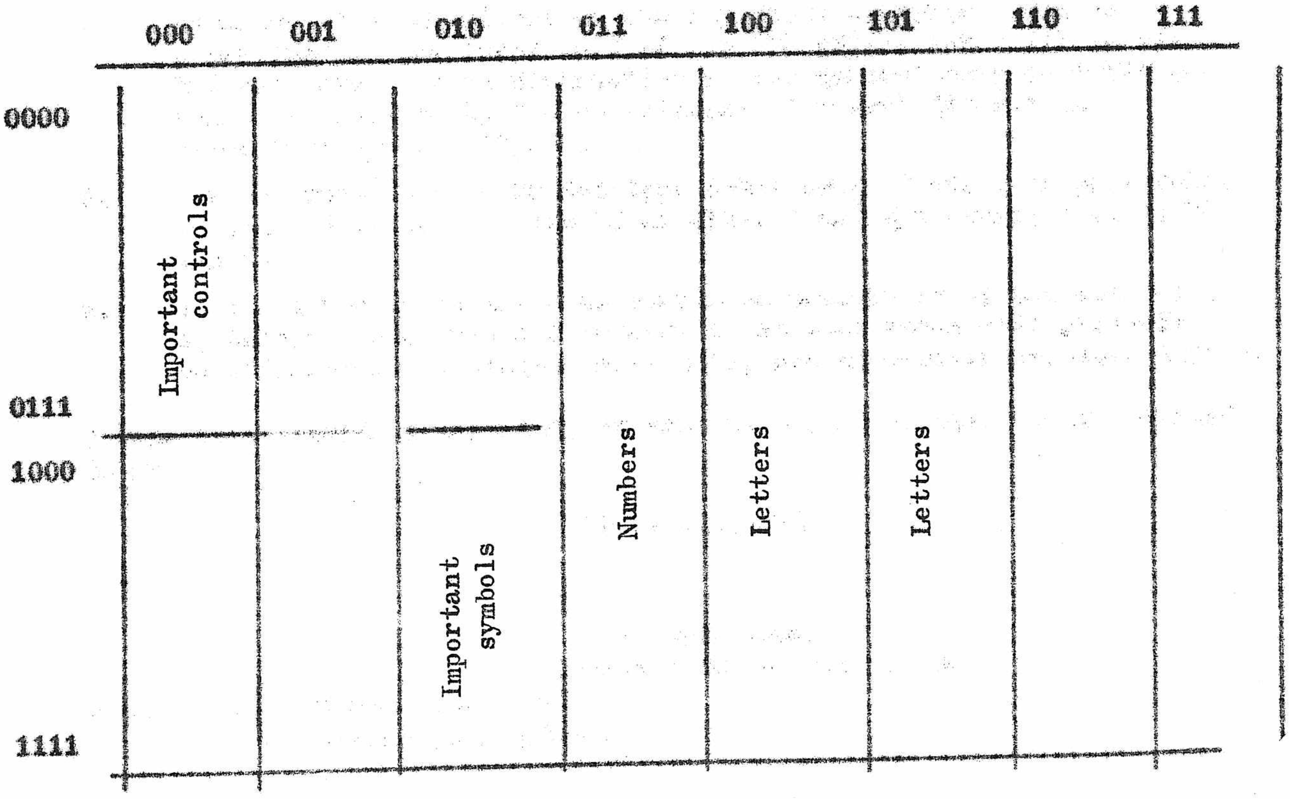
We recognised the advance represented by the 7-bit coding scheme, and the importance of the techniques used to transform it into smaller sets; we were appreciative that the numbers and letters follow the ISO code, and that allowance has been made for sterling numbers and the extra Continental European letters. We were glad to see the treatment of Null and Blank (which has been of some difficulty to us). We recognised the suitability of the scheme for 8-bit computers, for line-printers with 64 printing symbols, for transmission systems using 8 or 5-unit characters, and for typewriter keyboards.

We felt that only one significant difficulty remains. We could not see any simple transformation or contraction into a 6-bit set which is suitable for general-purpose computer use.

In Britain, and also Continental Europe, virtually all computers use6-bit characters. In order that a code may be usable with these computers, the control functions and data delimiters need to collate low to the letters, numbers and special symbols (for the same reasons as they have been arranged in this way in the ASA 7-bit system).

In the BSI and ECMA schemes, the first 8 positions of the 6-bit set have been allocated for this purpose, and the meanings of the various characters have been chosen so that they can be used to control machines, to control the format or layout of documents, or as data delimiters, as appropriate. This arrangement has been found satisfactory in practice, and is considered to be the optimum within a 6-bit set.

We saw that a simple transformation could be obtained between the ASA 7-bit scheme and a general-purpose 6-bit set if these important controls were grouped in the first half of the 000 "stick" of the ASA scheme, as shown generally:-



The contraction would be obtained by ignoring the 6th bit, whereupon the 000 and 010 sticks would overlap to give the required type of 00 stick in the 6-bit set.

Within the 000 half-stick the controls would follow the BSI code (or ECMA proposals) thus (using the Teletype nomenclature):-

	000
0000	Null
0001	CR
0010	LF
0011	V. Tab
0100	H. Tab
0101	
0110	LC
0111	UC

We recognised that on some machines, e.g. Teletypewriter, this would necessitate a change of the keys used to generate some of these combinations. We hoped, however, that this might be accepted to permit use of the code with European computers.

Three other points were mentioned, which we felt to be of less importance than that above.

1. An alternative treatment of the important controls might be to place them in the lower half of the 000 stick, but still in the BSI sequence. Then contraction to the general-purpose 6-bit set could be obtained by two operations: "invert" 4th bit in stick 000; ignore 6th bit.
2. For contraction to a Baudot-type 5-bit set, it might be preferable to code FS at 0101110 and LS at 0101111 and not equate them to UC and LC.
3. We did not wish to lay much stress on uniformity of the special symbols, because from one country to another there will probably be differences of usage, importance, and typewriter keyboard layout.

I hope this record of our study of the ASA proposals will be of interest to you.

Yours sincerely,

H. McG. Ross
Computer Sales Department

c.c. Mr. C.E. Mason, ASA X3.2
Mr. J.P. Richardson, B.S.A

BERKE
LIBRA
UNIVERSIT
CALIFOR

PLANNING A COMPUTER SYSTEM

PROJECT STRETCH

CONTRIBUTORS

Richard S. Ballance
Robert W. Bemer
Gerrit A. Blaauw
Erich Bloch
Frederick P. Brooks, Jr.
Werner Buchholz
Sullivan G. Campbell
John Cocke
Edgar F. Codd
Paul S. Herwitz
Harwood G. Kolsky
Edward S. Lowry
Elizabeth McDonough
James H. Pomerene
Casper A. Scalzi

Edited by
WERNER BUCHHOLZ
SYSTEMS CONSULTANT
CORPORATE STAFF, RESEARCH AND ENGINEERING
INTERNATIONAL BUSINESS MACHINES CORPORATION

New York Toronto London 1962
McGRAW-HILL BOOK COMPANY, INC.

Chapter 6

CHARACTER SET

by R. W. Bemer and W. Buchholz

6.1. Introduction

Among the input and output devices of a computer system, one can distinguish between those having built-in codes and those largely insensitive to code. Thus typewriters and printers necessarily have a fixed code that represents printable symbols to be read by the human eye; a code must be chosen for such a device in some more or less arbitrary fashion, and the device must make the transformation between code and symbol. Data storage and transmission devices, on the other hand, such as magnetic tape units and telephone transmission terminals, merely repeat the coded data given to them without interpretation, except that some code combinations may possibly be used to control the transmission process. (Strictly speaking, storage and transmission devices do generally limit the code structure in some respect, such as maximum byte size, so that code sensitivity is a matter of degree.)

For the inherently code-sensitive devices to be attached to a new computer system, an obvious choice of character set and code would have been one of the many sets already established. When the existing sets were reviewed, however, none were found to have enough of the system characteristics considered desirable. In fact, it became clear that about the only virtue of choosing an already established set is that the set exists. Accordingly, it was decided, instead, to devise a new character set expressly for use throughout a modern computer system, from input to output. The chief characteristic of this set is its extension to many more different characters than have been available in earlier sets. The extended set designed for the 7030 (Fig. 6.1) contains codes for 120 different characters, but there is room for later expansion to up to 256 characters including control characters. In addition, useful subsets have been defined, which contain some but not all of these 120 characters and which use the same codes for the selected characters without translation.

It should be noted that the 7030 computer is relatively insensitive to the specific choice of code, and any number of codes could be successfully used in the system. For any particular application a specialized character code might be found superior. In practice, however, a large computer installation must deal with a mixture of widely different applications, and the designers have to choose a single character set as a compromise among conflicting requirements.

Bits 4-5-6-7	Bits 0-1-2-3							
	0000	0001	0010	0011	0100	0101	0110	0111
0000	Blank	[	ε	c	k	s	0	8
0001	±	⊃	+	C	K	S	$_0$	$_8$
0010	→ T	]	$	d	l	t	1	9
0011	≠	°	=	D	L	T	$_1$	$_9$
0100	∧	←	*	e	m	u	2	.
0101	{	≡	(	E	M	U	$_2$	:
0110	↑ 1	¬	/	f	n	v	3	–
0111	} 4	√ '	)	F	N	V	$_3$	?
1000	∨	%	,	g	o	w	4	
1001	∀	\	;	G	Θ	W	$_4$	
1010	↓	◊	'	h	p	x	5	
1011	||	|	"	H	P	X	$_5$	
1100	>	#	a	i	q	y	6	
1101	≧	!	A	I	Q	Y	$_6$	
1110	<	@	b	j	r	z	7	
1111	≦	~	B	J	R	Z	$_7$	

Fig. 6.1. 120-character set.

The purpose of this chapter is to list major requirements of a character set and code, and to point out how these requirements may or may not be met by the specific set to be described.

6.2. Size of Set

Present IBM 48-character sets consist of

1. 10 decimal digits
2. 26 capital letters
3. 11 special characters
4. 1 blank

Other manufacturers have employed character sets of similar or somewhat larger size.

Because a single set of eleven special characters is not sufficient, there exist several choices of special characters as "standard options."

Since this 48-character set is often represented by a 6-bit code, it is natural to try to extend it to 63 characters and a blank, so as to exploit the full capacity of a 6-bit code.[1] Although the extra sixteen characters would indeed be very useful, this step was thought not to be far-reaching enough to justify development of the new equipment that it would require.

As a minimum, a new set should include also:

5. 26 lower-case letters
6. The more important punctuation symbols found on all office typewriters
7. Enough mathematical and logical symbols to satisfy the needs of such programming languages as ALGOL[2,3]

There is, of course, no definite upper limit on the number of characters. One could go to the Greek alphabet, various type fonts and sizes, etc., and reach numbers well into the thousands. As set size increases, however, cost and complexity of equipment go up and speed of printing goes down. The actual choice of 120 characters was a matter of judgment; it was decided that this increment over existing sets would be sufficiently large to justify a departure from present codes and would not include many characters of only marginal value.

6.3. Subsets

Two subsets of 89 and 49 characters were chosen for specific purposes. The 89-character set (Fig. 6.2) is aimed at typewriters, which, with 44

---

[1] H. S. Bright, Letter to the Editor, *Communs. ACM*, vol. 2, no. 5, pp. 6–9, May, 1959 (a 64-character alphabet proposal).

[2] A. J. Perlis and K. Samelson, Preliminary Report: International Algebraic Language, *Communs. ACM*, vol. 1, no. 12, December, 1958.

[3] Peter Naur (editor), Report on the Algorithmic Language ALGOL 60, *Communs. ACM*, vol. 3, no. 5, May, 1960.

character keys, a case shift, and a space bar, can readily handle 89 characters. This subset was considered important because input-output typewriters can already print 89 characters without modification, and 44-key keyboards are familiar to many people.

The 49-character subset (Fig. 6.3) is the conventional set of "commercial" characters in a code compatible with the full set.[1] This subset is aimed at the *chain printer* mechanism used with the 7030, which can readily print character sets of different sizes but prints the larger sets at a reduced speed. The 49-character subset permits high-volume printing at high speed in a compatible code on jobs (such as bill printing) where the extra characters of the full set may not be needed. It should be noted that the 49-character set is not entirely a subset of the 89-character set.

Other subsets are easily derived and may prove useful. For example, for purely numerical work, one may wish to construct a 13-character set consisting of the ten digits and the symbols . (point) and – (minus), together with a special blank.

6.4. Expansion of Set

Future expansion to a set larger than 120 can take place in two ways. One is to assign additional characters to presently unassigned codes; allowance should then be made for certain control codes which will be needed for communication and other devices and which are intended to occupy the high end of the code sequence.

The second way is to define a shift character for "escape" to another character set.[2] Thus, whenever the shift character is encountered, the next character (or group of characters) identifies a new character set, and subsequent codes are interpreted as belonging to that set. Another shift character in that set can be used to shift to a third set, which may again be the first set or a different set. Such additional sets would be defined only if and when there arose applications requiring them.

6.5. Code

In choosing a code structure, many alternatives were considered. These varied in the basic number of bits used (i.e., the byte size) and in the number of such bytes that might be used to represent a single (print-

---

[1] Note that this is one character larger than the previously referred-to 48-character set. The additional special character was introduced in 1959 on the printer of the IBM 1401 system; but its use has not become firmly established, partly because it has no counterpart on the keypunch. Thus the 48- and 49-character sets are, in effect, the same set.

[2] R. W. Bemer, A Proposal for Character Code Compatibility, *Communs. ACM*, vol. 3, no. 2, February, 1960.

Bits 4-5-6-7	Bits 0-1-2-3							
	0000	0001	0010	0011	0100	0101	0110	0111
0000	Blank		Ɛ	c	k	s	0	8
0001			+	C	K	S	*0*	*8*
0010			$	d	l	t	1	9
0011			=	D	L	T	*1*	*9*
0100			*	e	m	u	2	.
0101			(	E	M	U	*2*	:
0110			/	f	n	v	3	-
0111			)	F	N	V	*3*	?
1000			,	g	o	w	4	
1001			;	G	O	W	*4*	
1010			'	h	p	x	5	
1011			"	H	P	X	*5*	
1100			a	i	q	y	6	
1101			A	I	Q	Y	*6*	
1110			b	j	r	z	7	
1111			B	J	R	Z	*7*	

FIG. 6.2. 89-character set.

able) character. Among the alternatives were the following:

Single 6-bit byte with shift codes interspersed
Double 6-bit byte = single 12-bit byte[1]
Single 8-bit byte
Single 12-bit byte for "standard" characters (punched-card code) and two 12-bit bytes for other characters

Some of these codes represented attempts to retain partial compatibility with earlier codes so as to take advantage of existing equipment.

---

[1] R. W. Bemer, A Proposal for a Generalized Card Code for 256 Characters, *Communs. ACM*, vol. 2, no. 9, September, 1959.

Bits 4-5-6-7	Bits 0-1-2-3							
	0000	0001	0010	0011	0100	0101	0110	0111
0000	Blank		Ɛ				0	8
0001				C	K	S		
0010			$				1	9
0011				D	L	T		
0100			*				2	.
0101				E	M	U		
0110			/				3	-
0111				F	N	V		
1000		%	,				4	
1001				G	O	W		
1010		◊	'				5	
1011				H	P	X		
1100		#					6	
1101			A	I	Q	Y		
1110		@					7	
1111			B	J	R	Z		

FIG. 6.3. 49-character set.

These attempts were abandoned, in spite of some rather ingenious proposals, because the advantages of partial compatibility were not enough to offset the disadvantages.

The 8-bit byte was chosen for the following reasons:

1. Its full capacity of 256 characters was considered to be sufficient for the great majority of applications.
2. Within the limits of this capacity, a single character is represented by a single byte, so that the length of any particular record is not dependent on the coincidence of characters in that record.
3. 8-bit bytes are reasonably economical of storage space.

4. For purely numerical work, a decimal digit can be represented by only 4 bits, and two such 4-bit bytes can be packed in an 8-bit byte. ~~Although such packing~~ of numerical data is not essential, it is a common practice in order to increase speed and storage efficiency. Strictly speaking, 4-bit bytes belong to a different code, but the simplicity of the 4-and-8-bit scheme, as compared with a combination 4-and-6-bit scheme, for example, leads to simpler machine design and cleaner addressing logic.

omit

5. Byte sizes of 4 and 8 bits, being powers of 2, permit the computer designer to take advantage of powerful features of binary addressing and indexing to the bit level (see Chaps. 4 and 5).

The eight bits of the code are here numbered for identification from left to right as 0 (high-order bit) to 7 (low-order bit). "Bit 0" may be abbreviated to B_0, "bit 1" to B_1, etc.

6.6. Parity Bit

For transmitting data, a ninth bit is attached to each byte for parity checking, and it is chosen so as to provide an odd number of 1 bits. Assuming a 1 bit to correspond to the presence of a signal and assuming also an independent source of timing signals, *odd parity* permits all 256 combinations of 8 bits to be transmitted and to be positively distinguished from the absence of information. The parity bit is identified here as "bit P" or B_P.

The purpose of defining a parity bit in conjunction with a character set is to establish a standard for communicating *between* devices and media using this set. It is not intended to exclude the possibilities of error correction or other checking techniques *within* a given device or on a given medium when appropriate.

6.7. Sequence

High-equal-low comparisons are an important aspect of data processing. Thus, in addition to defining a standard code for each character, one must also define a standard comparing (*collating*) sequence. Obviously, the decimal digits must be sequenced from 0 to 9 in ascending order, and the alphabet from A to Z. Rather more arbitrary is the relationship between groups of characters, but the most prevalent convention for the 48 IBM "commercial" characters is, in order:

(Low) Blank
Special characters . ⌑ & $ * – / , % # @
Alphabetic characters A to Z
(High) Decimal digits 0 to 9

Fundamentally, the comparing sequence of characters should conform to the natural sequence of the binary integers formed by the bits of that code. Thus 0000 0100 should follow 0000 0011. Few existing codes have this property, and it is then necessary, in effect, to translate to a special internal code during alphanumeric comparisons. This takes extra equipment, extra time, or both. An important objective of the new character set was to obtain directly from the code, without translation, a usable comparing sequence.

A second objective was to preserve the existing convention for the above 48 characters within the new code. This objective has not been achieved because of conflicts with other objectives.

The 7030 set provides the following comparing sequence without any translation:

(Low) Blank
Special characters (see chart)
Alphabetic characters a A b B c C to z Z
Numerical digits 0 $_0$ 1 $_1$ to 9 $_9$
Special characters . : — ?
(High) Unassigned character codes

Note that the lower- and upper-case letters occur in pairs in adjacent positions, following the convention established for directories of names. (There appeared to be no real precedent for the relative position within the pair. The case shift is generally ignored in the sequence of names in telephone directories, even when the same name is spelled with either upper- or lower-case letters. This convention is not usable in general, since each character code must be considered unique.)

The difference between this comparing sequence and the earlier convention lies only in the special characters. Two of the previously available characters had to be placed at the high end, and the remaining special characters do not fall in quite the same sequence with respect to one another. It was felt that the new sequence would be quite usable and that it would be necessary only rarely to re-sort a file in the transition to the 7030 code. It is always possible to translate codes to obtain any other sequence, as one must do with most existing codes.

6.8. Blank

The code 0000 0000 is a natural assignment for the *blank* (i.e., the nonprint symbol that represents an empty character space). Not only should the *blank* compare lower than any printable character, but also absence of bits (other than the parity bit) corresponds to absence of mechanical movement in a print mechanism.

Blank differs, however, from a *null* character, such as the all-ones code found on paper tape. *Blank* exists as a definite character occupying a definite position on a printed line, in a record, or in a field to be compared.

A *null* may be used to delete an erroneous character, and it would be completely dropped from a record at the earliest opportunity. *Null*, therefore, occupies no definite position in a comparing sequence. A *null* has not been defined here, but it could be placed when needed among the control characters.

Considering numerical work only, it would be aesthetically pleasing to assign the all-zeros code to the digit zero, that is, to use 0000 as the common zone bits of the numeric digits (see below). In alphanumeric work, however, the comparing sequence for *blank* should take preference in the assignment of codes.

6.9. Decimal Digits

The most compact coding for decimal digits is a 4-bit code, and the natural choices for encoding 0 to 9 are the binary integers 0000 to 1001. As mentioned before, two such digits can be packed into an 8-bit byte; for example, the digits 28 in packed form could appear as

0010 1000

If decimal digits are to be represented unambiguously in conjunction with other characters, they must have a unique 8-bit representation. The obvious choice is to spread pairs of 4-bit bytes into separate 8-bit bytes and to insert a 4-bit prefix, or *zone*. For example, the digits 28 might be encoded as

zzzz 0010 *zzzz* 1000

where the actual value of each zone bit *z* is immaterial so long as the prefix is the same for all digits.

This requirement conflicted with requirements for the comparing sequence and for the case shift. As a result, the 4-bit byte is offset by 1 bit, and the actual code for 28 is

011**0** **01**00 011**1** **000**0

This compromise retains the binary integer codes 0000 to 1001 in adjacent bit positions, but not in either of the two positions where they appear in the packed format.

The upper-case counterparts of the normal decimal digits are assigned to italicized decimal subscripts.

6.10. Typewriter Keyboard

The most commonly found devices for key-recording input to a computer system are the IBM 24 and 26 keypunches, but their keyboards are not designed for keying both upper- and lower-case alphabetic characters. The shifted positions of some of the alphabetic characters are used to punch numerical digits. For key-recording character sets with much more than the basic 48 characters, it is necessary to adopt a keyboard convention different from that of the keypunch. The 89-character subset was established to bring the most important characters of the full set within the scope of the common typewriter, thus taking advantage of the widespread familiarity with the typewriter keyboard and capitalizing on existing touch-typing skills as much as possible.

The common typewriter keyboard consists of up to 44 keys and a separate case-shift key. To preserve this relationship in the code, the 44 keys are represented by 6 bits of the code (B_1 to B_6) and the case shift by a separate bit (B_7). The case shift was assigned to the lowest-order bit, so as to give the desired sequence between lower- and upper-case letters.

For ease of typing, the most commonly used characters should appear in the lower shift ($B_7 = 0$). This includes the decimal digits and, when both upper- and lower-case letters are used in ordinary text, the lower-case letters. (This convention differs from the convention for single-case typewriters presently used in many data-processing systems; when no lower-case letters are available, the digits are naturally placed in the same shift as the upper-case letters.) It is recognized that the typewriter keyboard is not the most efficient alphanumeric keyboard possible, but it would be unrealistic to expect a change in the foreseeable future. For purely numerical data, it is always possible to use a 10-key keyboard either instead of the typewriter keyboard or in addition to it.

It was not practical to retain the upper- and lower-case relationships of punctuation and other special characters commonly found on typewriter keyboards. There is no single convention anyway, and typists are already accustomed to finding differences in this area.

6.11. Adjacency

The 52 characters of the upper- and lower-case alphabets occupy 52 consecutive code positions without gaps. For the reasons given above, it was necessary to spread the ten decimal digits into every other one of twenty adjacent code positions, but the remaining ten positions are filled with logically related decimal subscripts. The alphabet and digit blocks are also contiguous. Empty positions for additional data and control characters are all consolidated at the high end of the code chart.

This grouping of related characters into solid blocks of codes, without empty slots that would sooner or later be filled with miscellaneous characters, assists greatly in the analysis and classification of data for editing purposes. Orderly expansion is provided for in advance.

6.12. Uniqueness

A basic principle underlying the choice of this set is to have only one code for each character and only one character for each code.

Much of the lack of standardization in existing character sets arises from the need for more characters than there are code positions available in the keying and printing equipment. Thus, in the existing 6-bit IBM character codes, the code 001100 may stand for any one of the three characters @ (at), – (minus), and ' (apostrophe). The 7030 set was required to contain all these characters with a unique code for each.

The opposite problem exists too. Thus, in one of the existing 6-bit codes, – may be represented by either 100000 or 001100. Such an embarrassment of riches presents a logical problem when the two codes have in fact the same meaning and can be used interchangeably. No amount of comparing and sorting will bring like items together until one code is replaced by the other everywhere.

In going to a reasonably large set, it was necessary to resist a strong temptation to duplicate some characters in different code positions so as to provide equal facilities in various subsets. Instead, every character has been chosen so as to be typographically distinguishable if it stands by itself without context. Thus, for programming purposes, it is possible to represent any code to which a character has been assigned by its unique graphic symbol, even when the bit grouping does not have the ordinary meaning of that character (e.g., in operation codes).

In many instances, however, it is possible to find a substitute character close enough to a desired character to represent it in a more restricted subset or for other purposes. For example, = (equals) may stand for ← (is replaced by) in an 89-character subset. Or again, if a hyphen is desired that compares lower than the alphabet, the symbol ~ (a modified tilde) is preferred to the more conventional – (minus).

A long-standing source of confusion has been the distinction between upper-case "oh" (O) and zero (0). Some groups have solved this problem by writing zero as Ø. Unfortunately, other groups have chosen to write "oh" as Ø. Neither solution is typographically attractive. Instead, it is proposed to modify the upper-case "oh" by a center dot (leaving the zero without the dot) and to write and print "oh" as ʘ whenever a distinction is desired.

Various typographic devices are used to distinguish letters (I, l, V, etc.) from other characters [| (stroke), 1 (one), ∨ (or), etc.]. It is suggested that the italicized subscripts be underlined when handwritten by themselves, for example, $_{\underline{5}}$.

6.13. Signs

The principle of uniqueness implies a separate 8-bit byte to represent a plus or a minus sign. Keying and printing equipment also require separate sign characters. This practice is, of course, rather expensive in storage space, but it was considered superior to the ambiguity of present 6-bit codes where otherwise "unused" zone bits in numerical fields are used to encode signs. If the objective is to save space, one may as well abandon the alphanumeric code quite frankly and switch to a 4-bit decimal coding with a 4-bit sign digit, or go to the even more compact binary radix.

6.14. Tape-recording Convention

As has been remarked before, data-recording media such as magnetic tape and punched cards are not inherently code-sensitive. It is obviously necessary, though, to adopt a fixed convention for recording a code on a given medium if that medium is to be used for communication between different systems.

Magnetic tape with eight, or a multiple of eight, information tracks permits a direct assignment of the 8 bits in the 7030 code to specific tracks. Magnetic tape with six information tracks requires some form of *byte conversion* to adapt the 8-bit code to the 6-bit tape format. The convention chosen is to distribute three successive 8-bit bytes over four successive 6-bit bytes on tape. This convention uses the tape at full efficiency, leaving no gaps except possibly in the last 6-bit byte, which may contain 2 or 4 nonsignificant 0 bits, depending on the length of the record.

Thus successive 8-bit bytes, each with bits B_0 to B_7, are recorded as shown in Table 6.1.

TABLE 6.1. CONVENTION FOR RECORDING 8-BIT CODE ON 6-TRACK TAPE

Track	*Bits*					
0	B_0	B_6	B_4	B_2	B_0	
1	B_1	B_7	B_5	B_3	B_1	
2	B_2	B_0	B_6	B_4	B_2	
3	B_3	B_1	B_7	B_5	B_3	etc.
4	B_4	B_2	B_0	B_6	B_4	
5	B_5	B_3	B_1	B_7	B_5	

The parity bit is not shown. The parity bits for the 6-bit tape format are, of course, different from those of the 8-bit code; so parity conversion must be provided also.

6.15. Card-punching Convention

Since 80-column punched cards are a common input medium, a card-punching convention for the 120 characters is likewise desirable. After the possibility of a separate card code for the 120 characters was considered—a code having the conventional IBM card code as a subset[1]—

---

[1] *Ibid.*

it was concluded that it would be better to punch the 8-bit code directly on the card. This does not preclude also punching the conventional code (limited to 48 characters) on part of the card for use with conventional equipment. Code translation is then needed only whenever the conventional card code is used; otherwise translation would be required for every column if advantage is to be taken of the new code in the rest of the system.

The punching convention is given in Table 6.2.

In addition, both hole 12 and hole 11 are to be punched in column 1 of every card containing the 7030 code, besides a regular 7030 character, so as to distinguish a 7030 card from cards punched with the conventional code. Eight-bit punching always starts in column 1 and extends as far as desired; a control code END (0 1111 1110) has been defined to terminate the 8-bit code area. Conventional card-code punching should be confined to the right end of those cards identified with 12-11 punching in column 1.

TABLE 6.2. CONVENTION FOR PUNCHING 8-BIT CODE ON CARDS

Card row	*Bit*
12	—
11	—
0	—
1	B_P
2	B_0
3	B_1
4	B_2
5	B_3
6	B_4
7	B_5
8	B_6
9	B_7

Since the parity bit is also punched, the 7030 area of a card contains a checkable code. Note that "blank" columns in this area still have a hole in the B_P row. If only part of the card is to be punched, however, it is possible to leave the remaining columns on the right unpunched.

6.16. List of 7030 Character Set

A list of the 7030 character-set codes and graphic symbols is shown for reference in Fig. 6.4, which includes the names of the characters.

Code P 0123 4567	Character	Name
1 0000 0000		Blank (Space)
0 0000 0001	±	Plus or minus
0 0000 0010	→	Right arrow (Replaces)
1 0000 0011	≠	Not equal
0 0000 0100	∧	And
1 0000 0101	{	Left brace
1 0000 0110	↑	Up arrow (Start superscript)
0 0000 0111	}	Right brace
0 0000 1000	∨	Or (inclusive)
1 0000 1001	∀	Exclusive Or
1 0000 1010	↓	Down arrow (End superscript)
0 0000 1011	‖	Double lines
1 0000 1100	>	Greater than
0 0000 1101	≥	Greater than or equal
0 0000 1110	<	Less than
1 0000 1111	≤	Less than or equal
0 0001 0000	[	Left bracket
1 0001 0001	⊃	Implies
1 0001 0010	]	Right bracket
0 0001 0011	°	Degree
1 0001 0100	←	Left arrow (Is replaced by)
0 0001 0101	≡	Identical
0 0001 0110	¬	Not
1 0001 0111	√	Square root (Check mark)
1 0001 1000	%	Percent sign
0 0001 1001	\	Left slant (Reverse divide)
0 0001 1010	◊	Lozenge (Diamond) (Note)
1 0001 1011	|	Absolute value (Vertical line)
0 0001 1100	#	Number sign
1 0001 1101	!	Exclamation point (Factorial)
1 0001 1110	@	At sign
0 0001 1111	~	Tilde (Hyphen)
0 0010 0000	&	Ampersand
1 0010 0001	+	Plus sign
1 0010 0010	$	Dollar sign
0 0010 0011	=	Equals
1 0010 0100	*	Asterisk (Multiply)
0 0010 0101	(	Left parenthesis
0 0010 0110	/	Right slant (Divide)
1 0010 0111	)	Right parenthesis
1 0010 1000	,	Comma
0 0010 1001	;	Semicolon
0 0010 1010	'	Apostrophe (Single quote)
1 0010 1011	"	Ditto (Double quote)
0 0010 1100	a	
1 0010 1101	A	
1 0010 1110	b	
0 0010 1111	B	
1 0011 0000	c	
0 0011 0001	C	
0 0011 0010	d	
1 0011 0011	D	
0 0011 0100	e	
1 0011 0101	E	
1 0011 0110	f	
0 0011 0111	F	
0 0011 1000	g	
1 0011 1001	G	
1 0011 1010	h	
0 0011 1011	H	
1 0011 1100	i	
0 0011 1101	I	
0 0011 1110	j	
1 0011 1111	J	

Note: The character ⌑ has also been used.

FIG. 6.4. List of 7030 codes and characters. (*Continued on next page.*)

Code P 0123 4567	Character	Name
0 0100 0000	k	
1 0100 0001	K	
1 0100 0010	l	
0 0100 0011	L	
1 0100 0100	m	
0 0100 0101	M	
0 0100 0110	n	
1 0100 0111	N	
1 0100 1000	o	
0 0100 1001	O	
0 0100 1010	p	
1 0100 1011	P	
0 0100 1100	q	
1 0100 1101	Q	
1 0100 1110	r	
0 0100 1111	R	
1 0101 0000	s	
0 0101 0001	S	
0 0101 0010	t	
1 0101 0011	T	
0 0101 0100	u	
1 0101 0101	U	
1 0101 0110	v	
0 0101 0111	V	
0 0101 1000	w	
1 0101 1001	W	
1 0101 1010	x	
0 0101 1011	X	
1 0101 1100	y	
0 0101 1101	Y	
0 0101 1110	z	
1 0101 1111	Z	
1 0110 0000	0	Zero
0 0110 0001	$_0$	Subscript zero
0 0110 0010	1	One
1 0110 0011	$_1$	Subscript one
0 0110 0100	2	Two
1 0110 0101	$_2$	Subscript two
1 0110 0110	3	Three
0 0110 0111	$_3$	Subscript three
0 0110 1000	4	Four
1 0110 1001	$_4$	Subscript four
1 0110 1010	5	Five
0 0110 1011	$_5$	Subscript five
1 0110 1100	6	Six
0 0110 1101	$_6$	Subscript six
0 0110 1110	7	Seven
1 0110 1111	$_7$	Subscript seven
0 0111 0000	8	Eight
1 0111 0001	$_8$	Subscript eight
1 0111 0010	9	Nine
0 0111 0011	$_9$	Subscript nine
1 0111 0100	.	Period (point)
0 0111 0101	:	Colon
0 0111 0110	–	Minus sign
1 0111 0111	?	Question mark

FIG. 6.4 (*Continued*)

Chapter 7

VARIABLE-FIELD-LENGTH OPERATION

by G. A. Blaauw, F. P. Brooks, Jr., and W. Buchholz

7.1. Introduction

Chapter 4 dealt with the fact that natural data units for fixed-point-arithmetic, logical, and editing operations vary considerably in length and structure. The variable-field-length instructions of the 7030 have been designed to make it possible to specify these natural data units simply and directly, thus saving time, space, and programming effort.

The variable-field-length (VFL) data-handling operations may be divided into three classes: (1) arithmetical, (2) radix-conversion, and (3) logical-connective operations. VFL arithmetical and logical-connective operations are both used also for processing alphanumeric data.

The VFL instructions include the basic arithmetical repertoire (LOAD, STORE, ADD, COMPARE, MULTIPLY, DIVIDE) as well as interesting new operations and features. More important, however, is the method of data definition employed by all VFL instructions. Each field, regardless of length, is treated as a separate entity independent of its neighbors. Each numerical field may have its own sign, if a sign is desired. Any overflow beyond the end of the specified field is signaled, but the next adjacent field is protected from inadvertent carry propagation. Similarly, any loss of significant bits caused by storing a result in a field of limited size is signaled. A *result zero* indicator shows the state of only the desired field, no more and no less.

The flexibility needed for VFL operations is achieved most economically by a serial data-handling mechanism. Serial data handling is relatively slow, but the objective here is not high speed for individual instructions. (Where arithmetical speed is of the essence, the unnormalized floating-point mode should be used for fixed-point arithmetic—see Chap. 8.) The VFL instructions are intended for such operations on complex data structures as format conversion and arranging for printing. Such operations can be performed by a serial VFL unit faster than by

Berne

AGENDA

EUROPEAN TRIP ON CODE ALPHABET

JANUARY

3, 4 and 5 PARIS; Conferences with
J. Birle', Director General, AFNOR
(AFNOR is Alliance Francaise de Normalization - equivalent to ASA)
(AFNOR has the Secretariate of the subgroup of ISO-TC97 working on code alphabet)

Dreyfus
R. Dreifuss, ECMA Committee on Code Alphabet

H. Feissel, Co. de Machines Bull

J. Krieger, IATA

8,9,10 and 11 LONDON; Conferences with
H. M. Ross, Ferranti, Ltd. (English code expert)

K. Amos, Creed and Co.

J. Rhodes, British Post Office
(Chairman, United Kingdom Delegation to CCITT)

B. Roston, NCR Co.
(Chairman, BSI Committee on Paper Tape Standards)

W. Eadie, IBM, Standards Coordinator

15 MILAN; Conference with
Dr. Zannini, Olivetti Computers
(Chairman, Italian Code Alphabet Task Group reporting to ISO)

16, 17 MUNICH; Conferences with
H. Wusteney, Siemens-Halske Co.
Dr. Lockeman

18, 19 STUTTGART; Conference with
Dr. Kaiser, Lorenz and Co.

22 — 10:00 AM Sat 20th GENEVA; Conference with Geneva 32-67-10
J. Besseyre, Senior Counsel, CCITT
(per recommendation of V. N. VAUGHAN)

23 TURIN; Conference with
A. Sarti, Olivetti Teleprinters

NMAH 310/2

PROPOSED AMERICAN STANDARD CODE FOR INFORMATION INTERCHANGE

UNOFFICIAL

January 11, 1962

NMAH 310/2

Proposed American Standard Code
for Information Interchange

1. Scope

1.1 This standard establishes a 7-bit coded character set for the interchange of data in an information processing system. No

1.2 This standard does not define the means by which the coded set is to be recorded in any physical medium. The standard code does not include any redundancy or define techniques for error control. Further, it does not specify a standard collating sequence. No

2. Standard Code

See table and notes below

			b7	0	0	0	0	1	1	1	1
			b6	0	0	1	1	0	0	1	1
			b5	0	1	0	1	0	1	0	1
b4	b3	b2	b1								
0	0	0	0	NULL	R1 ON	b̸	0	!	P	UNASSIGNED	UNASSIGNED
0	0	0	1	SOM	R2 ON	'	1	A	Q		
0	0	1	0	M OFF	X ON	"	2	B	R		
0	0	1	1	RU	EOA	<	3	C	S		
0	1	0	0	WRU	R1 OF	>	4	D	T		
0	1	0	1	EOT	R2 OF	$	5	E	U		
0	1	1	0	H TAB	X OFF	%	6	F	V		
0	1	1	1	FF	EOM	&	7	G	W		
1	0	0	0	V TAB	DD	(	8	H	X		
1	0	0	1	①	DD	)	9	I	Y		
1	0	1	0	BELL	DD	.	↑	J	Z		
1	0	1	1	①	DD	,	←	K	@		
1	1	0	0	CR	DD	*	+	L	\		ACK
1	1	0	1	LF	DD	:	;	M	[		①
1	1	1	0	UC	DD	—	=	N	]		ESC
1	1	1	1	LC	DD	/	?	O	#		DEL

GRAPHICS

① - Unassigned Controls
DD - Reserved for Data Delimiter

Z @ O \ 30°

Note 1: Deviations from the standard may create serious difficulties in information interchange and should be used only with full cognizance of the parties involved.

Note 2: Unassigned codes are reserved for future standardization. Their use in information interchange prior to such standardization is a deviation from the standard.

January 26, 1962

PROPOSAL FOR COMPATIBLE CODE FOR INFORMATION INTERCHANGE

FERRANTI LTD
ELECTRICAL & GENERAL ENGINEERS

21 PORTLAND PLACE
LONDON W·1

TELEGRAMS
DIGIT, WESDO
LONDON

TELEPHONE
LANGHAM 9211

HEAD OFFICE & WORKS
HOLLINWOOD, LANCS
TELEGRAMS
FERRANTI, HOLLINWOOD
TELEPHONE
FAILSWORTH 2000

YOUR REF.
OUR REF. 441/HMcGR/CTJ

29th January, 1962

2/5

Mr. J.F. Auwaeter,
Director of Product Development,
Teletype Corporation,
SKOKIE, ILLINOIS
U. S. A.

Dear Mr. Auwaeter:

- CODING -

Following our very valuable discussion when you visited this country, I have attempted to think out a compatible code which might meet the needs of all groups working in this field. I attach a copy of this proposal.

I have attempted to find a way of grouping the controls together in an orderly way, taking into account the needs for contraction and expansion of the codes and for a number of different machines and systems. I have tried to meet the points you made about simplicity in the design of your new Teletypewriters and associated equipment. I hope it will not be too difficult to make the "control" key invert bit-4 as well as forcing bit-7 = 0. Although the layout of the keyboard may not be quite as nice as yours, I hope you may feel it tolerable in the interests of general compatibility. The controls in the bottom half of 001-zone has been arranged in pairs, because this makes it easier to accommodate FS and LS, and makes things tidier if any one of these machines is not required; however, these could be re-arranged if you particularly desired.

I hope you will feel these proposals offer a prospect of progress, and I look forward to any further discussion we may have about them.

Yours sincerely,

H. McG. Ross.

H. McG. Ross
Computer Sales Department

Encl:

NMAH
310/2

January 26, 1962

PROPOSAL FOR COMPATIBLE CODE FOR INFORMATION INTERCHANGE

1. This proposal for a compatible code for information interchange attempts to meet all the requirements of the ASA, ECMA, BSI and EEA Coding Committees. In particular, it combines all the features considered best in the schemes so far developed by these groups.

2. In seeking compatibility, the main task is to attain uniformity of those factors which:-

(i) Affect the design of equipment;

(ii) Affect programming conventions, and especially input routines;

(iii) Affect the operational planning of complete systems.

The factors which have the greatest importance from these points of view are the controls or functions. Consequently, attention in the attached chart is concentrated on these. The coding of numbers and letters has already been agreed; it is considered that the choice and coding of special symbols is less important and can be tackled later.

3. The attached chart is based on the ASA X3.2 proposed code, but shows only the first 2 zones containing the controls; the remainder of the Code Table would follow the ASA proposals, but possibly with amendment of the special symbols.

4. In addition to the type of controls specified in the ASA scheme, it is also shown how these proposals could be applied to a number of other systems.

5. In designing this code, the controls have been arranged in the following 4 groups:

(i) Upper half of 000-zone contains the most-important and most-used controls;

(ii) Upper half of 001-zone contains the data delimiters;

(iii) Lower half of 000-zone contains controls which are primarily required to control various systems; in a way, these can be regarded as characters which are not part of the true information in the message;

(iv) Lower half of 001-zone contains controls which turn off and on auxiliary equipment.

6. In order to obtain an orderly and logical arrangement of the code, the controls have been arranged within the first 5 positions of each of the first 3 of the above half-zones in a sequence which corresponds to that of Table 1 of the report of the ECMA Task Group, November 8-9. Thus, controls which are equivalent to one another are on the same "level" or "sub-level".

7. Care has been taken to fit in with a typewriter-style keyboard and to meet inexpensively the needs of Teletypewriters and Flexowriters. The keyboard layout obtained is shown in the second figure. On Teletypewriters and Flexowriters, all the control codes can be made from the letter-keys in conjunction with a "control" (or "auxiliary") key, which inhibits the emission of bit-7 and also inverts bit-4 (bits are numbered bit-7.....bit-1). Only Motor Off requires special treatment, i.e. bit-4 should not be inverted. WRU and Bell are placed fairly near to their current positions on the CCITT keyboard. Keys M, and also perhaps N and 0, are free to be paired up with special symbols coded at the foot of the 101-zone. And keys X,Y and Z are available for making the 3

CBI 67/
1/6 ?

8. The following contractions from the 7-bit set are available:-

(i) Dense 6-bit sub-set obtained from zones 010-110, as in ASA proposals;

(ii) Dense 4-bit graphic sub-set, as in ASA proposals;

(iii) All the transformations in Mr. Bemer's note of 12th September, 1961, (Table 1), are unaltered;

(iv) A 6-bit set of the BSI/ECMA type, suitable for general-purpose use with European computers, is easily obtained by ignoring bit-6 and moving bit-7 into its place;

(v) 5-bit sub-set is easily obtained with FS and LS mapping over Escape and Delete.

9. As an alternative arrangement, the upper and lower halves of zones 000 and 001 could be transposed. Then the control key on Teletypewriters would not need to invert bit-4. In order to contract to the type of 6-bit sub-set required for general-purpose computer use, bit-4 in these 2 zones would then have to be inverted before the contraction of 8(iv) is carried out. Although this would be easy enough for computers to do, it might prove more difficult for simpler machines. As a consequence, and in case designers of future equipment did not fully understand the situation, there might, therefore, be a tendency for two 6-bit forms of code to come into use (one with the controls in the two halves of these zones transposed, and the other without). It is considered this would be unfortunate, and it would also present a serious difficulty in placing the more-important symbols in zone 010.

10. The characters for Special Shift and Normal Shift are probably not required in any 7-bit system, but they are essential for the general-purpose 6-bit computer code of 8(iv), and they are needed to control the conversion back from the 6-bit to 7-bit sets. Their places are, therefore, shown in the 7-bit Table because these must not be used for anything else. (Normal Shift is that which contains the numbers; in accordance with ordinary data processing and punched card practice, present users of the BS code also place capital letters in Normal Shift, but some users in Continental Europe may prefer to have lower-case letters in Normal Shift).

11. It is assumed that data delimiters in ASA terminology are, at least in part, equivalent to the separators of information in ECMA terminology.

12. Special points concerning ASA telecommunication system. This is slanted towards the needs of telecommunications (a group of 3 characters), and of message-switching systems (store and re-transmit) (a group of 4 characters), and the remote starting and stopping of auxiliary machines - the codes for the latter are arranged in pairs which makes it tidier if one such machine is not required or is to be replaced by another, and it makes it easier to fit in FS and LS.

13. Special points concerning British Standard Code. This proposal requires the following changes from the BS Code:-

(i) Transposition of Tabulate and an unallocated control combination. This affects machines and programming techniques, and is an example of the critical consequences resulting from a change of a control character. It is 2½ years since some British organisations began to use codes of the BS type and they may now be unable to follow this transposition; others now using the BS code may be less committed, and able to change.

3/

(ii) Backspace and CR are treated as alternates, this being felt to be justified because existing machines use one or the other, not both, and CR may become less important in future (see also 14.(ii).

(iii) CR is recoded; no-one is yet known to be using the existing BS code for this.

(iv) The 8-track punched tape code would be changed to follow the zone-structure of the ASA scheme.

14. Special points concerning Flexowriter system:-

(i) Newline means combined CR and LF; throughout this note the latter are used in their CCITT meaning, not in the confusing Flexowriter meaning.

(ii) The machine is designed so that it is not damaged if a Backspace character is received immediately after a Newline character.

(iii) X = Readers
R_1 = Printer
R_2 = Punch 1
FS = Red, this trick having been found useful for compatibility between certain single-shift and double-shift machines.

15. The Magnetic Tape and Machine Tool Control Systems show possible extensions of the scheme.

H. McG. Ross

Ferranti Ltd.,
London Computer Centre,
21 Portland Place,
LONDON, W1

HMcGR/CTJ

PROPOSAL FOR COMPATIBLE CODE FOR INFORMATION INTERCHANGE

This chart shows only the controls in the first two zones. The remainder of the codes would follow the ASA X3.2 proposals (except perhaps for some re-arrangement of special symbols).

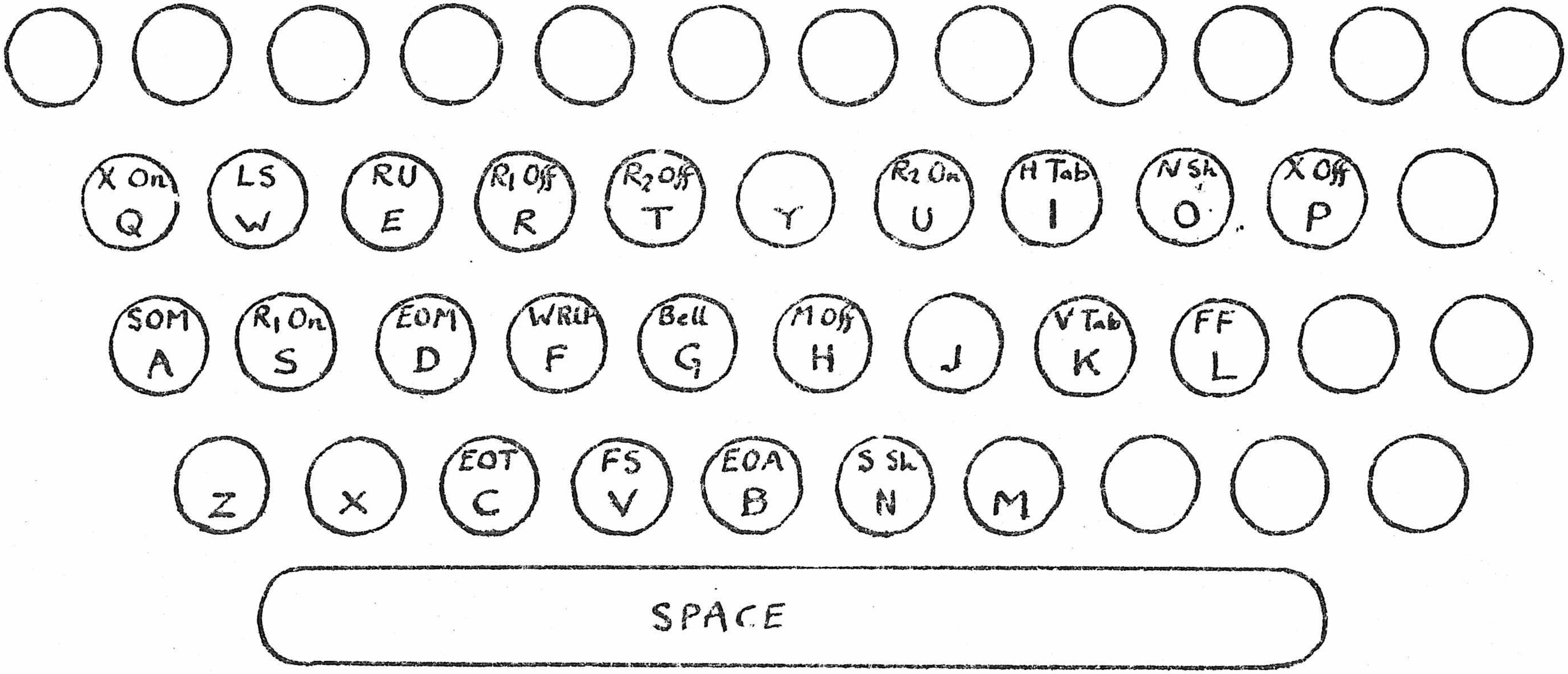
	ASA Telecommunication System		BSI/ECMA Notation		Flexowriter System		Magnetic Tape System	Machine Tool Control System
	000	001	000	001	000	001	000	000
0000	Null	Delimiter 0	No Action	End of Sub-field	Blank Tape		Filler	No Action
0001	H Tab	Delimiter 1	H Tab	End of Field	H Tab			
0010	LF	Delimiter 2	Newline	End of Item	Newline (CR)			
0011	V Tab	Delimiter 3	Paperthrow	End of Group	Line Finder			
0100	FF	Delimiter 4	End of Form	End of Job				
0101	CR		Backspace or CR		Backspace			
0110	Special Shift		Shift Out		LC Shift			
0111	Normal Shift and TapeFeed		Shift In and Run Out		UC Shift and Tape Feed			
1000	M Off	X Off				Switch Readers (Control)	Tape Mark	End of Address Symbol
1001	SOM	X On			Skip	Punch 2 On	Word Separator	End of Coordinate Word
1010	EOA	R_1 Off				Non Print	Block Mark	End of Point
1011	EOT	R_1 On				Print Restore	Group Mark	End of Operation
1100	EOM	R_2 Off			Stop	Punches Off	Record Mark	End of Job
1101	RU	R_2 On	ARQ			Punch 1 On	Delta	Reverse
1110	WRU	FS		FS		Red or Switch Colors		
1111	Bell	LS		LS		Black		

In all systems, Space (blank) = 0100000, Escape = 1111110, Delete = 1111111.

In ASA Telecommunication system, ACK = 1111100.

H. McG. R.

January 25, 1962



PAIRING OF CONTROLS ON TYPEWRITER-STYLE KEYBOARD

Jan 26 1962

File
X3.2

January 11, 1962

TO: X3.2 Members and Alternates

FROM: X3.2 Secretary

The Minutes of the December 11-14, 1961 meeting were mailed without the Summary page, which is enclosed. Please attach this page to your copy of the minutes, and then be prepared to discuss and approve them at the January 31 session.

The November 8-10, 1961 Meeting Minutes have not been distributed pending a careful review by Mr. Macon. They are particularly important in the approval procedure of a Draft Standard and are being scrutinized accordingly. They will, however, be out prior to the January 31 session to enable their approval at that date, and submission to BEMA and X3.

R E Utman

R. E. Utman

mw

The National Cash Register Company
ELECTRONICS DIVISION
1401 EAST EL SEGUNDO BOULEVARD • HAWTHORNE, CALIFORNIA • PLymouth 7-1811

30 January 1962

2/2

Mr. John Auwaerter
Teletype Corporation
5555 Touhy Avenue
Skokie, Illinois

Dear John:

This is the trip report which I have distributed to the X3.2 Committee, to interested parties within NCR and to the people we visited on our trip. I covered most heavily the areas directly concerned with the code and touched only lightly on the general telecommunication discussions. I hope that I covered the trip adequately and accurately. I am sure that a supplementary report by you would be more than welcome by all interested parties, should you wish to write one.

Hope you enjoyed the camel ride. Give my best to Irene and I will be seeing you in New York probably, shortly.

Sincerely,

Leon Bloom

Leon Bloom

LB:jo

Enclosure

NMAH 36/2

REPORT ON DISCUSSIONS BETWEEN
John Auwaerter and Leon Bloom
of ASA X3.2
AND MEMBERS OF EUROPEAN DATA PROCESSING COMMUNITY

SUBJECT:
INTERNATIONAL STANDARDIZATION
OF CHARACTER CODES AND SETS

25 January 1962

NMAH 310/2

INTRODUCTION

We will attempt to describe the long complex trip just completed by summarizing first the major purposes, impressions and accomplishments of the trip and then describing, in some detail, each of our visits. It will, of course, be impossible to convey in a report of reasonable length all the information gathered on the trip. It is our hope that we will have the opportunity to discuss the trip in more detail, with the X3.2 Committee as well as with others concerned with the subject.

SUMMARY

The purpose of our visit to Europe was to meet informally with those people in the data processing and telecommunications industries in Europe who were active in the standardization of character codes and sets. It was our intention to explain in some detail the thinking which caused the proposed American Standard to develop as it did, the status of the proposed standard and the plans that various companies in the United States, particularly the Bell System, had for implementation of the new code. We then hoped to answer any questions which may have been generated either by documents which had been received before our visit or by the discussions during our visit.

At the same time, we felt it would be useful to get further information on the progress of standardization work in Europe and to get a better understanding of the factors influencing this work.

In all respects our visit was highly successful. We visited a majority of the people in Europe who are active in the standardization of character sets and codes. At each stop we found a tremendous willingness to cooperate in an exchange of information. (As a matter of fact, it was quite unanimous that it would have been desirable to have begun a cooperative effort about 6 months ago so that each of us could have had the benefit of the others needs and thoughts.) In many instances we found a lack of complete understanding of the motives and intentions in certain critical areas of the work being done on both sides. We believe that any misunderstandings have now been eliminated so that further exchanges will be on a sounder basis.

The most important result of our visit, however, was the generation of a plan, agreed to by all those we visited, by which an attempt would be made to reach a compromise agreement on a proposal to be submitted to the ISO TC/97 meeting to be held in France sometime in October 1962. It was highly gratifying to see the degree of enthusiasm which developed over such a possibility. It is our hope and expectation that the same kind of enthusiasm will be generated in the United States and that every effort will be made to reach an international agreement at the earliest possible date. It is our firm conviction that such an agreement is highly desirable and is achievable. Further, it is our feeling that if agreement is not achieved in the immediate future, each group will have proceeded in its own direction to the extent that an international standard will not be achievable for many years to come.

FRANCE:

Persons Visited	Company Affiliation	Standardization Organization
M. Giret	AFNOR	AFNOR
M. Weill	BULL - Head Logical Standards	
P. Dreyfus	BULL	Vice-Chairman ECMA
M. Durand	BULL	ECMA TC - 1
M. Pedretti	IBM	Treasurer ECMA
M. Feissel	BULL	Chairman WGB ISO TC - 97

The visit to France covered three main areas:

1. Development of a plan for activation of WGB ISO TC - 97.
2. Standardization activities in France.
3. An introduction to the activities of ECMA TC - 1 on character sets.

Plan for WGB

Mr. Feissel reported that only 3 groups at the most were expected to have sent proposed standards to WGB in time for the October meeting of ISO TC - 97. These groups were the BSI and the ASA for certain, and ECMA TC - 1 as a possibility. It was agreed that it would be highly desirable if a single proposal from all three groups could be developed in time for the TC - 97 meeting so as to avoid further solidifying each of their positions along different lines. In order to make a serious attempt at accomplishing this the following plan was proposed:

1. Mr. Auwaerter and Mr. Bloom would, during their visit to Europe, attempt to determine whether such a combined proposal was achievable within the available time. Mr. Weill planned to meet the American representatives in Geneva at the conclusion of their trip to discuss the conclusions. (The conclusions were positive.)

2. Assuming a positive conclusion X3.2 would send delegates to the ECMA TC - 1 meeting in March in an attempt to develop such a proposal.
3. In the meanwhile, a meeting of WGB would be called for May 1962.
4. During the May meeting a draft proposal to ISO TC - 97 would be developed.
5. The time from May to October would be required for reproduction, transmission, translation and study of the proposal by various members of ISO TC - 97. (A minimum of four months is required for consideration of such proposals.)

Standardization Activities in France

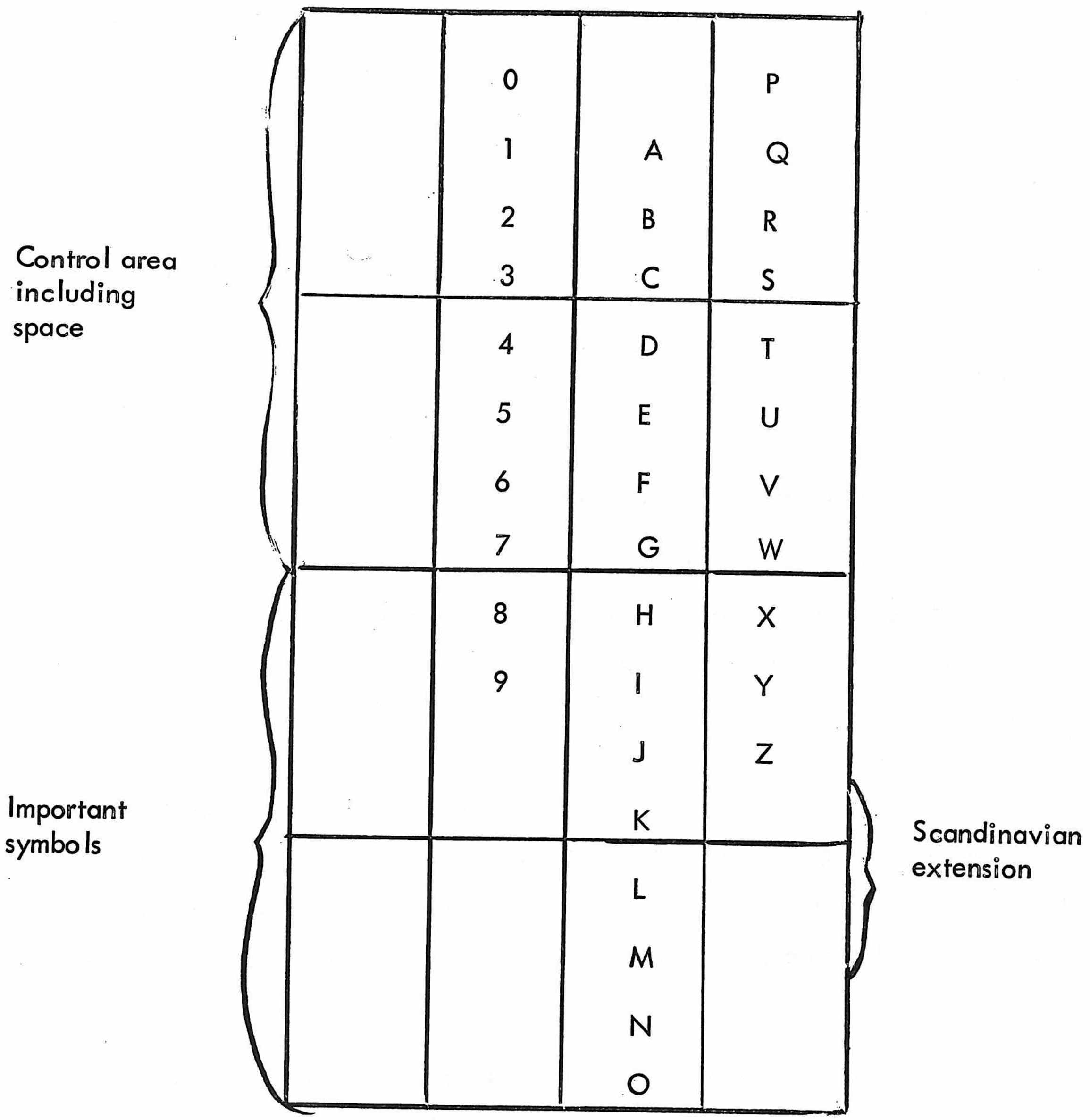
At least six multi-lettered standardization groups active in fields related to data processing were described. The function and interrelation of these groups is too complex to explain (or even understand fully). Suffice it to say, France is very active in standardization work and is completely serious about efforts to arrive at workable standards in the data processing and related fields.

Activities of ECMA TC - 1

This group began its work early in 1961 with the same general scope for Europe as X3.2 has for the United States. (ECMA is an organization composed of representatives from eighteen European computer manufacturers.) Thus far about six meetings have been held and the work has progressed to the point where:

1. General agreement has been reached that a 6-bit code will be standardized.
2. The location of the alphabet, numerals, escape, delete and the three extra alphabetic characters required by the Scandinavians has been agreed to.
3. The area in which important controls are to be located has been agreed to.
4. The area in which the most important special symbols are to be located has been agreed to.
5. The exact controls and symbols to be included as well as the sequence of these controls and symbols has not been settled.
6. At their last meeting, following receipt of information about the work of X3.2 on a 7-bit set, several proposals were advanced on possible methods of expansion to a 7-bit set.

Areas of Agreement by ECMA TC - 1



	0		P
	1	A	Q
	2	B	R
	3	C	S
	4	D	T
	5	E	U
	6	F	V
	7	G	W
	8	H	X
	9	I	Y
		J	Z
		K	
		L	
		M	
		N	
		O	

FIGURE I

It should be noted that this work closely resembles the work of the British Standards Institute so that there is complete agreement between the X3.2 proposal, the BSI standard and ECMA TC - 1 on the location of the numerals and the alphabet. When the X3.2 7-bit set is collapsed into a 6-bit set, it becomes obvious that the basic area of conflict between the European work and that of the Americans is primarily in the location of the important controls and secondarily in the location and sequence of the important symbols. The reasons for the conflict and the philosophy behind the placement and sequence of the control became clear later in our trip.

ENGLAND

Persons Visited	Company Affiliation	Standardization Organization
K. Amos	British Post Office	CCITT
Holland-Martin		President of ECMA
H. McG. Ross	Ferranti Ltd.	BSI, ECMA TC - 1
Mr. Chadwick	BSI	BSI
T. Cluff	ICT	BSI, ECMA TC - 1
J. Smith		BSI
Mr. Crawley	ITT	BSI, CCITT
A. Curnow	NCR	BSI, ECMA TC - 1

The visit to England covered two main points:

1. The status of the BSI code.
2. The differences and possible areas of compromise between the BSI code, the ECMA work and the X3.2 proposed standard.

Status of the BSI Code

The BSI code has just been published and is therefore now an approved standard. A number of companies in England have been working on implementation of the standard and several companies have already included a small number of features of the code in their equipment. The BSI has been working on the development of this standard for three years.

It was generally conceded, however, that should international agreement be reached on a standard, although different in some respects from that of the BSI, the British Standards Institute would wholeheartedly cooperate in the establishment of such a standard. It was felt that establishment of such a standard would be highly desirable.

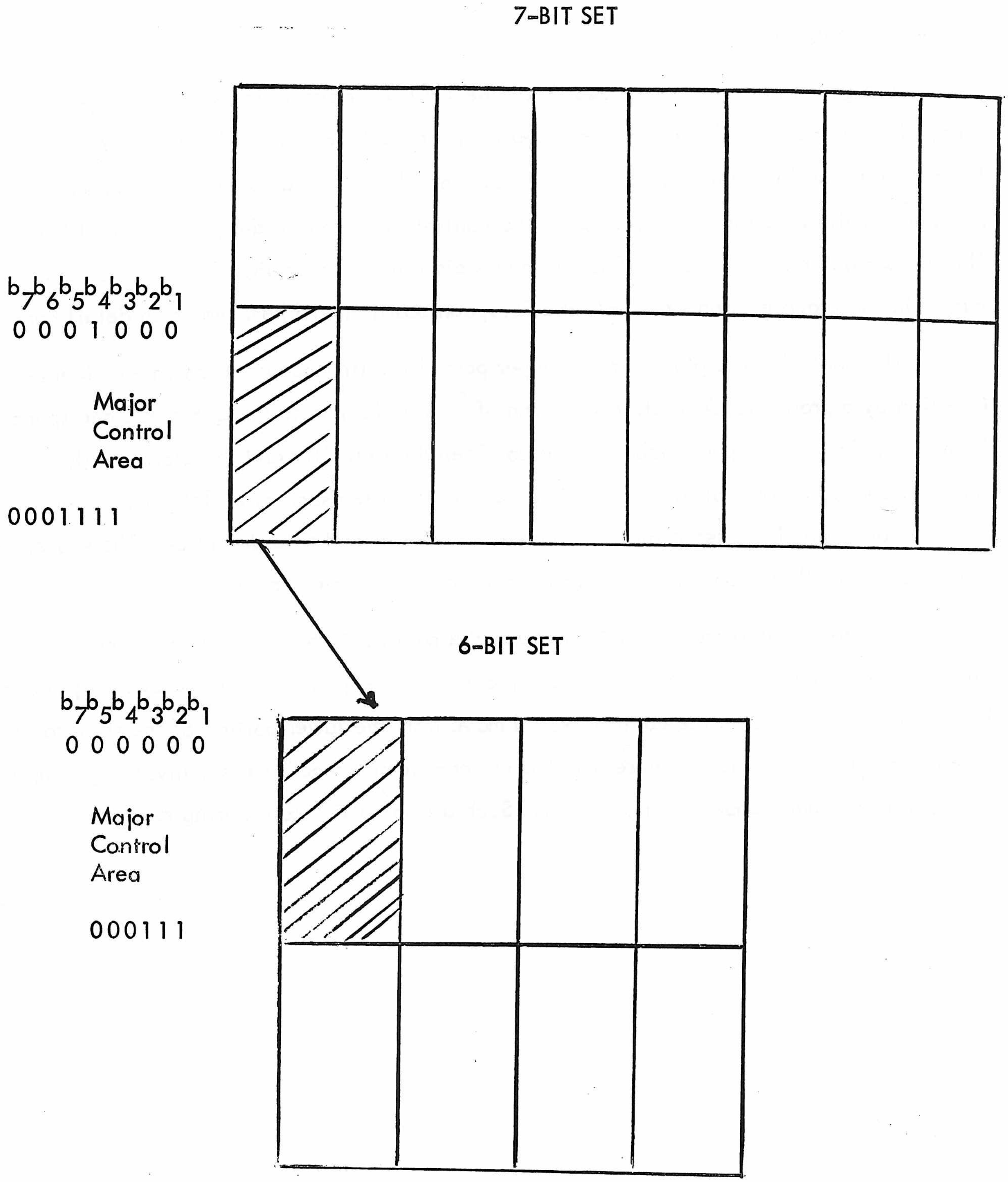
Areas of Compromise Between Codes

It was generally agreed that the USA had gone too far down the path of the 7-bit set and the Europeans too far in working toward a 6-bit set for there to be any serious hope of either changing their basic organization. (This feeling was prevalent among the data processing people we visited but a number of the communicators showed definite interest in eventually going to a 7 or 8 level set.)

The goal which it was decided should be strived for and which showed the possibility of being met was that of having a 6-bit set for the Europeans which bore a simple logical relationship to the American 7-bit set. The alphabet and numerals presented no problem but the location and sequence of the controls and symbols did present a problem. The reasons behind each choice of location was discussed. Since the X3.2 position is discussed in the proposed standard, only the European position will be covered in detail here.

The controls were placed in the upper part of the first quadrant for ease of identification by a program. A single comparison of $\leq$ would separate these controls and space from data. For this purpose space may be considered a delimiter and therefore, in this sense, a control. Some of the controls included in this area such as H Tab, Line Feed, Vert Tab and Form Feed serve the dual purpose of control and data delimiter. The sequence and grouping of these controls therefore becomes of major significance.

If each side persists in its wish to keep the controls in the sections in which they are currently proposed, a possible means of compromise is to expand or collapse by manipulating bit b_4 so that on collapse the controls would move from the lower portion of the stick to the upper portion. If this method were adopted the compromise would merely involve getting the controls in the same sequence in both sets. Such a compromise could bring results.



Expansion or contraction - Manipulation of Control Area

7 $\longrightarrow$ 6: eliminate b_6, invert b_4

6 $\longrightarrow$ 7: $b_6 = b_7$, invert b_4

ITALY

Persons Visited	Company Affiliation	Standardization Organization
G. Sacerdoti	Ollivetti	ECMA - TC - 1
R. Gottlieb	Ollivetti	Formerly with Major Leubbert at Ft. Monmouth
L. Zannini	Ollivetti	Chairman WGD ISO TC - 97

Our discussions here centered on the work of ECMA TC - 1. The need for standardization on an international level was strongly backed but so were the concepts of the 6-bit code as they were being developed by ECMA. It was recognized, however, that to develop a 6-bit code without consideration of expansion, or a 7-bit code without considering the collapse, would be foolish. The outlined methods of expansion and collapse seemed reasonable as a starting point for further discussion and it was indicated that every effort should be made to reach agreement according to the plan worked out in France and along the lines worked out in England.

GERMANY - MUNICH

People Visited	Company Affiliation	Standardization Organization
Dr. Lockemann	Seamans	Chairman ECMA TC - 1
H. Wusteney	Seamans	
H. Hennig	Seamans	

The discussions here centered about the needs of the telecommunicators in Europe and how these needs relate to the data processors. The TELEX network is the telecommunication network used throughout Europe (about 100,000 machines are in use). They make use of a shifting 5-level code and there is no evidence that there will be a change from this in the immediate future. (There was some expression of opinion that a change might be forthcoming in five to ten years and that the change might incorporate a higher level code. Any change in Europe would probably be gradual however, since most teleprinters in the TELEX network are customer owned, making for a greater reluctance to obsolete a piece of equipment.)

Another aspect of the situation which was discussed at each of the countries visited and which was discussed at some length here, was that of keyboard layout and the significance assigned to this factor in establishing the code. It was pointed out, and we had also concluded from our own observations, that keyboards were far less standard in Europe than they were in the United States. England was the only country where : (colon) and ; (semicolon) were paired on a key and the pairings on other keys showed no standardization and very little in the way of a pattern. (In some instances even the alphabet placement was not standard with as many as 6 letters being placed differently than on United States keyboards.) Since there was no generally accepted keyboard layout, little emphasis was placed on arranging the code to facilitate an historic keyboard arrangement.

Dr. Lockemann, as Chairman of ECMA TC - 1 expressed wholehearted approval of our efforts to arrive at an agreement by May so that a single proposal for an international standard might be submitted to the October meeting of ISO TC - 97. The invitation to X3.2 to participate in the March meeting of ECMA TC - 1 was repeated. (We have since learned that a February meeting of TC - 1 has been called partially to consider the effects of the X3.2 proposed standard.)

GERMANY - Stutgart and Pfortsheim

People Visited	Company Affiliation	Standardization Organization
Dr. Kaiser	Standard Elektrik Lorenz	
H. Augustin	Standard Elektrik Lorenz	
Dr. Neubauer	Standard Elektrik Lorenz	ECMA TC - 1

During this visit we discussed further the needs of the European telecommunicators and clarified some aspects of the X3.2 work which were not clearly understood. In addition, however, Dr. Neubauer presented to us a revised approach to the use of controls in combination with data delimiters as part of the European 6-bit code. The approach will be proposed at the next meeting of ECMA TC - 1.

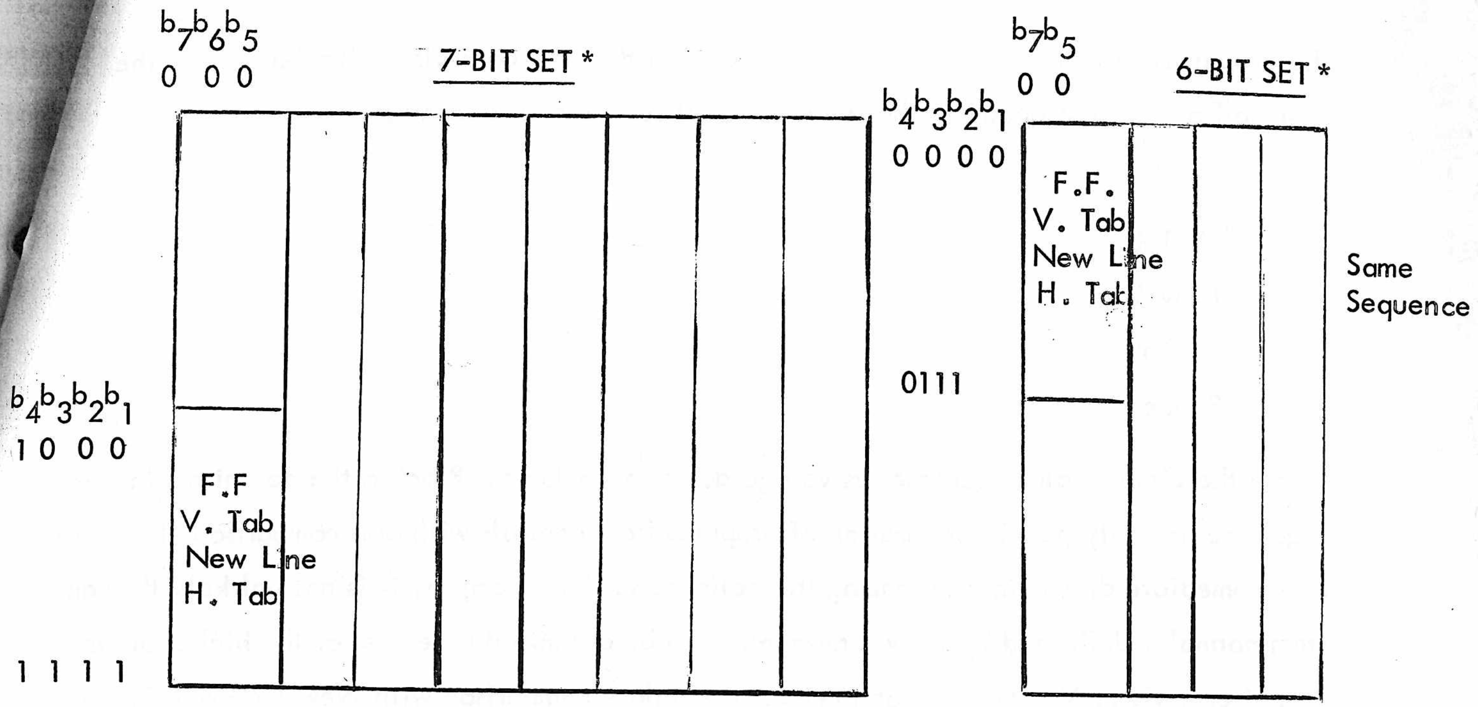
The system is based upon the idea that the lowest order delimiter should be closest to the graphics. This means that since it is proposed that the controls (delimiters) be in the first half of the first quadrant the sequence of these controls becomes:

FF
V. Tab
New Line
H. Tab
Space

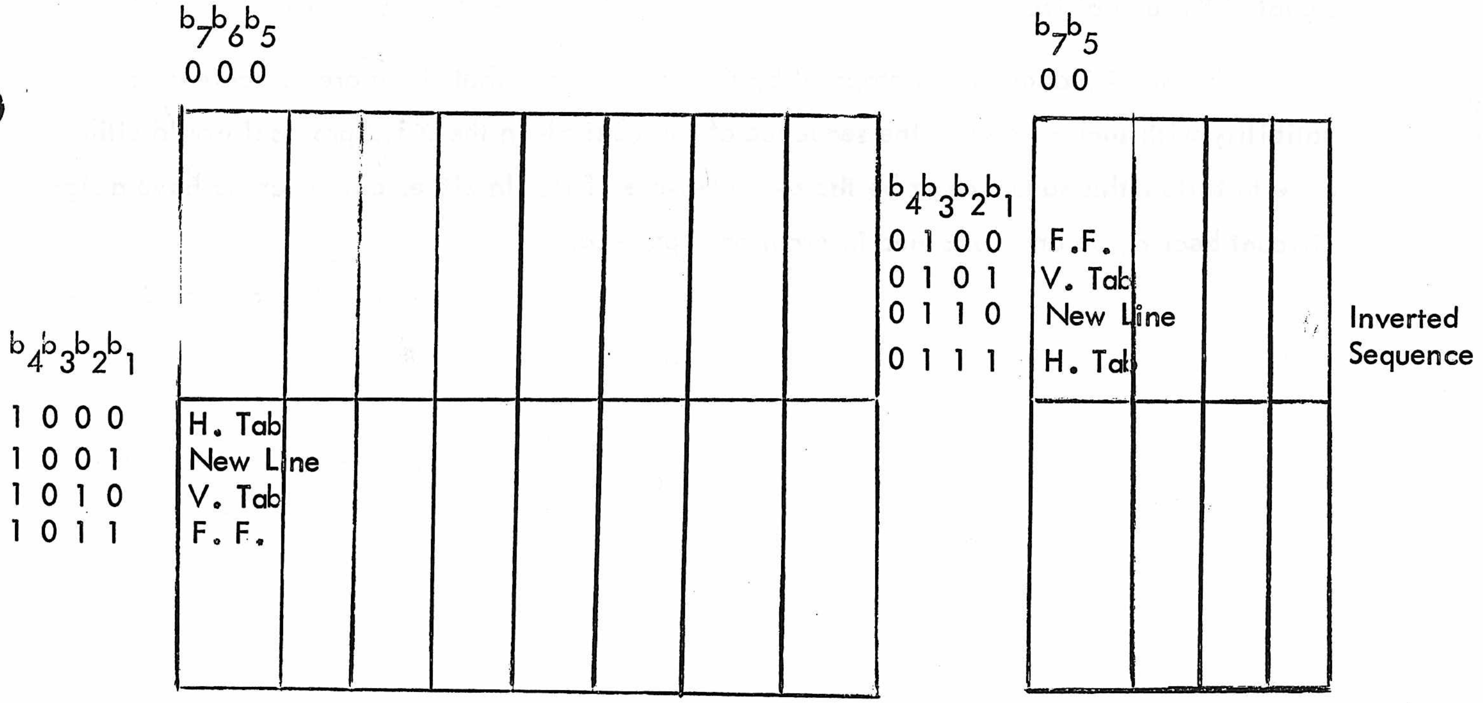
where the binary value increases as you go down the column. Placing the delimiters in this sequence not only permits separation of graphics from controls with one comparison, but permits immediate discrimination among the delimiters. For example, it is not unlikely that an item normally delimited by New Line might also be delimited by either of the higher order delimiters, viz., V. Tab and Form Feed. A single $\leq$ comparison with New Line will immediately determine the end of item regardless of which of the three possible delimiters actually signaled the end of item.

The significance of the proposal by Dr. Neubauer is that if we are to achieve compatibility with such a system, the sequence of the controls in the X3.2 proposal would either have to follow this sequence or be the exact reverse of it. In either case it could have a significant bearing on any combined international proposal.

Illustration of Conversion of Control Portion of 7-Bit Set to 6-Bit Set



7 ⟶ 6 Ignore b_6, Invert b_4



7 ⟶ 6 Ignore b_6, Invert b_4 b_3 b_2 b_1

* These diagrams ignore the placement of "space".

SWITZERLAND

Persons Visited	Company Affiliation	Standardization Organization
M. Bessayre	Senior Counsellor CCITT	CCITT
M. Weill	BULL	Representing AFNOR

In our meeting with M. Bessayre we discussed the activities of CCITT in their standardization activities for European telecommunication. It was pointed out that although there is current standardization on a 2-level 5-bit code, there was agitation on the part of the Czechs and Poles as long as three years ago for a study on extension of the alphabet. At that time, apparently, the United States and Great Britain indicated that they saw no need for an extension of the alphabet and as a result the study effort was dropped.

M. Bessayre has now received a communication from the American Telephone and Telegraph Company (written by John Booth, X3.2 member) which describes the X3.2 proposed standard. This document is being reproduced and will be distributed to CCITT members as an American proposal. It was M. Bessayre's belief that as European telecommunicators began switching from their present 50 baud system to 100 or 200 bauds, an expansion of the character set would accompany the change.

In our meeting with M. Weill we re-confirmed our conviction that a compromise between the European position and the American position could be reached if the plan originally outlined at the beginning of our trip were followed. It was further agreed that if agreement could not be reached at this time, it would be many years before a single international standard could be developed. We therefore strongly recommended that the proposed plan be put into effect immediately.